Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/003715

International filing date: 08 April 2005 (08.04.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 102004018608.1
Filing date: 16 April 2004 (16.04.2004)

Date of receipt at the International Bureau: 23 May 2005 (23.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

EP05/03715

BUNDESREPUBLIK DEUTSCHLAND

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 10 2004 018 608.1

Anmeldetag: 16. April 2004

Anmelder/Inhaber: Aventis Pharma Deutschland GmbH,
65929 Frankfurt/DE

Bezeichnung: Dual-transfected cells lines as in vitro screening tools for pharmaceutical compound profiling: A model for heptobiliary elimination

IPC: C 12 N 5/06

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 8. März 2005
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Schäfer

A 9161
03/00
EDV-L

Dual-transfected cells lines as in vitro screening tools for pharmaceutical compound profiling: A model for heptobiliary elimination

A cell line transfected with both human NTCP and human MRP2 suitable as an in vitro tool for pharmaceutical compound profiling: A model for hepatobiliary elimination

The invention is referring to several dual-transfected cell lines expressing human NTCP (Na/taurocholate Cotransporting Protein; SLC10A1) together with human BSEP (Bile Salt Export Pump; ABCB11) or human MRP2 (Multidrug Resistance Protein; ABCC2) suitable as an in vitro tool for pharmaceutical compound profiling particularly as a model for hepatobiliary elimination.

Transcellular transport processes across biological barriers are playing an essential role in the absorption, distribution and elimination of drugs and xenobiotics. In intestinal epithelial cells transport proteins are involved in the absorption or exclusion of drugs into/from the systemic circulation. Transport proteins expressed in the cells of the blood-brain-barrier are involved in the excision of toxic (and also therapeutic) compounds from the brain. In kidney proximal tubule epithelia cells and in hepatocytes transport proteins are involved in the elimination of various endogenous and exogenous substances. In the hepatocyte this elimination process is mediated by (i) uptake from blood into the cell via the basolateral membrane and (ii) elimination into the bile via the apical membrane. Paracellular diffusion in all biological barriers is anticipated by the formation of tight junctions (molecular structures providing a very tight contact between adjacent cells) separating the apical from the basolateral membrane compartment. Cell lines that have maintained the features of a biological barrier (development of basolateral and apical membrane compartments separated by tight junctions) are also called polarized cell lines. Usually these cell lines grow in a single cell layer designated as monolayer.

The basolateral (also called sinusoidal) uptake of compounds into the human hepatocyte is mediated by a variety of different transport proteins: Bulky organic anions are taken up by OATP8 (SLC21A8), OATP2 (SLC21A6, also known as OATP-C or LST1) and OATP-B (SLC21A9) whereas bile acids and possibly other compounds similar to bile acids are accepted by the Sodium Taurocholate Co-transporting Polypeptide NTCP (SLC10A1). OAT2 (SLC22A7) imports small organic anions into the hepatocyte and OCT1 (SLC22A1) has been reported to facilitate the cellular uptake of small organic cations. At the apical (also called canalicular) membrane, the elimination of organic anions and bile salts into bile is mediated by the Multidrug-resistance Related Protein MRP2 (ABCC2) and the Bile Salt Export Pump BSEP (ABCB11), both members of the ABC transporter family (ATP-dependent export pumps). Other ABC Transporters also known to be expressed at the canalicular membrane of the human hepatocyte are the Breast Cancer Resistance Protein BCRP (ABCG2, also known as MXR) and the Multidrug-resistance protein MDR1 (ABCB1, also known as P-glycoprotein). These two proteins are thought to participate in the canalicular secretion process of amphiphilic and cationic compounds.

The multidrug resistance protein (MDR) family mediates the ATP-dependent unidirectional transport of conjugates of lipophilic substances with glutathione, glucuronate, or sulfate (Review: König J. et al., Biochimica et Biophysica Acta 1461 (1999) 377-394).

The bile salt pool undergoes an enterohepatic circulation that is regulated by distinct bile salt transport proteins, including the canalicular bile salt export pump BSEP, the ileal NA^+ -depentent bile salt transporter ISBT, and the hepatic sinusoidal Na^+ -taurocholate cotransporting polypeptide NTCP. Other bile salt transporters include the organic anion transporting polypeptides OATPs and the multidrug resistance-associatd proteins 2 and 3 MRP2,2 (Review: Trauner et al., Physiol. Rev. 83: 633-671,2003).

To date the only way to investigate hepatobiliary elimination of drugs is the bile fistula study done in rats, a complex and time-consuming *in vivo* assay. In addition the bile fistula study doesn't give any information about the molecular basis of the

elimination process. Here we describe a cell-based assay system in which a liver uptake transporter (in this case human NTCP) is constitutively expressed together with a liver export pump (usually an ABC transporter, in this case human MRP2 or human BSEP) in the polarized canine kidney cell line MDCKII (for reference see Louvard, 1980; Fuller et al., 1984). The resulting cell lines therefore express two transgenes in one cell line (i) human NTCP together with human BSEP (MDCKII-hNTCP/hBSEP) and (ii) human NTCP together with human MRP2 (MDCKII-hNTCP-hMRP2). The cells are cultivated in 6-well filter inserts on a porous filter membrane thereby separating a basolateral from an apical compartment (Fig.1). Putative substrates (e.g. pharmaceutical compounds) are added either to the apical or to the basolateral compartment for transcellular (vectorial) transport measurements in both directions. Samples are taken from each compartment after the indicated incubation periods. Substrates of a given transporter combination (in this case human NTCP and human BSEP or human NTCP and human MRP2) display a significant net transport from the basolateral to the apical compartment when compared to un-transfected or mono-transfected cells (Fig. 4 + 5).

With the aid of the described dual-transfected cell lines one will be able to predict whether the transporter combinations human NTCP / human BSEP or human NTCP / human MRP2 participates in the hepatobiliary elimination process of a pharmaceutical compound or not.

During pharmaceutical drug development compounds frequently fail to proceed into clinical development due to their weak pharmacokinetic profile. One reason for this is their rapid elimination via the liver into the bile. Examples for this phenomenon are the chlorogenic acid derivatives, potent and specific inhibitors of the glucose-6-phosphate translocase (a potential novel pharmaceutical target in patients with type II diabetes). Pharmacodynamic studies showed that the blood glucose lowering effect of several chlorogenic acid derivatives lasts only a short time after bolus i.v. injection in Wistar rats. Bile fistula studies indicated that this weak pharmacodynamic effect is the result of unsuitable pharmacokinetics due to a rapid hepatobiliary elimination process. With the aid of dual-transfected transporter cell lines one will be able to identify the transporters involved in the elimination process. In a next step the

compounds can than be chemically modified in a way that alleviates their affinity to the eliminating transporter to improve the compounds pharmacology.

Another application area would be the examination of potential drug-drug interactions at the level of transporters in the liver. For example competition of two drugs for one transporter in the liver may alter the pharmacokinetic profile of both compounds significantly. In addition, inhibition or activation of a transporter by one drug may also change the pharmacological behaviour of the transporters substrates.

The invention pertains to a mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hBSEP protein.

The first side of such a mammalian cell is e.g. the basolateral side and the second side is the apical side or the first side is e.g. the apical side and the second side is e.g. the basolateral side.

The mammalian cell can be chosen from epithelial cells in particular of the kidney, of the bowels systems, of the liver or of the blood/brain barrier.

Such cells can be immortalized cells or recombinant cells. The cells can be taken from ordinary mammalian tissue or a primary cell culture. Such cells can be also the cells of a cell culture as e.g. LLC-PK1 cells or MDCKII cells in particular also when carrying recombinant vectors.
The LLC-PK1 cells may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hBSEP protein (as exemplified by SEQ ID NO. 5). As alternative the hNTCP and the hBSEP proteins may be expressed simultaneously from the same vector construct. The MDCKII cell may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for

expressing the hBSEP protein (as exemplified by SEQ ID NO. 5). Such a mammalian cell of a MDCKII type carrying two separate vectors with ability of expressing hNTCP protein and hBSEP protein has been deposited with the DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig. The identification number is DSM ACC2643.

The invention pertains further to the manufacturing of a mammalian cell carrying a functional human NTCP protein and a functional human BSEP protein wherein

a] a mammalian cell is provided;
b] a vector is provided encompassing the coding sequence of hNTCP (e.g. vector as drafted according or equal to SEQ ID NO. 4);
c] a vector is provided encompassing the coding sequence of hBSEP (e.g. vector as drafted according or equal to SEQ ID NO. 5);
d] the mammalian cell from a] is transformed by a vector from b] and by a vector from c] either simultaneously or consecutively;
e] a double transfectant cell from d] is identified and propagated.

The mammalian cells that shall be provided for such a manufacturing may be epithelial cells in particular of the kidney, of the lowels system, of the liver or of the blood/brain barrier. Such cells can be also immortalized cells or recombinant cells. The cells can be taken from ordinary mammalian tissue or a primary cell culture. Such double transfectant cells according to e] resulting from the manufacturing as mentioned above may be e.g. LLC-PK1 cells or MDCKII cells in particular when carrying recombinant vectors. Such LLC-PK1 cells may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hBSEP protein (as exemplified by SEQ ID NO. 5). The hNTCP and the hBSEP proteins may be expressed also simultaneously from the same vector construct. Such MDCKII cells may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hBSEP (as exemplified by SEQ ID NO. 5). Such a mammalian cell of a MDCKII type carrying two separate vectors with ability of expressing hNTCP protein and hBSEP protein has been deposited with

the DSMZ-Deutsche Sammlung von Mikroorganismen and Zellkulturen GmbH, Mascheroder Weg 1 b, D-38124 Braunschweig. The identification number is DSM ACC2643.

The invention pertains further to a monolayer of cells comprising at least two cells of mammalian cells having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hBSEP protein. Such a monolayer could occupy a part or the whole of a solid surface. The invention pertains also to such a solid surface carrying a monolayer of such cells.

The solid surface could be formed by a plastic. The solid surface could also be part of a petri dish or a filter-insert.

The invention pertains further to a petri dish carrying a monolayer of cells as described before and as well as to a filter-insert carrying a monolayer of cells as described before. The membrane support of such a filter-insert could be made of polycarbonat and/or polyester. The pore size of the membrane support could be of 0.4 µm +/- 0.2 µm. An according example of a filter-insert has been depicted in Fig. 1. There are several manufacturers of filter-inserts that could be used for placing the said monolayer of cells on it. Such manufacturers are for example “Corning Inc., Corning, NY” or “Millipore”.

The mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hBSEP protein could be used e.g. for determining pharmacological profiles with respect to hepatobiliary elimination and/or renal excretion and/or brain resorption and/or intestinal resorption. For such use the mammalian cells could form part of a monolayer on a solid surface and/or a petri disk and/or on a filter insert.

The invention pertains further to a mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hMRP2 protein.

The first side of such a mammalian cell is e.g. the basolateral side and the second side is the apical side or the first side is e.g. the apical side and the second side is e.g. the basolateral side.

The mammalian cell can be chosen from epithelial cells in particular of the kidney, of the bowels system, of the liver or of the blood/brain barrier.

Such cells can be immortalized cells or recombinant cells. The cells can be taken from ordinary mammalian tissue or a primary cell culture. Such cells can be also the cells of a cell culture as e.g. LLC-PK1 cells or MDCKII cells particularly when carrying recombinant vectors.

The LLC-PK1 cells may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hMRP2 protein (as exemplified by SEQ ID NO. 6). As alternative the hNTCP and the hMRP2 proteins may be expressed simultaneously from the same vector construct. The MDCKII cell shall harbor a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hMRP2 protein (as exemplified by SEQ ID NO. 6). Such a mammalian cell of a MDCKII type carrying two separate vectors with ability of expressing hNTCP protein and hMRP2 protein has been deposited with the DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1 b, D-38124 Braunschweig. The identification number is DSM ACC2644.

The invention pertains further to the manufacturing of a mammalian cell carrying a functional human NTCP protein and a functional human MRP2 protein wherein

a] a mammalian cell is provided;

b] a vector is provided encompassing the coding sequence of hNTCP (e.g. a vector as drafted according or equal to SEQ ID NO. 4);

c] a vector is provided encompassing the coding sequence of hMRP2 (e.g. vector as drafted according or equal to SEQ ID NO. 6);

d] the mammalian cell from a] is transformed by a vector from b] and by a vector from c] either simultaneously or consecutively;

e] a double transfectant cell from d] is identified and propagated.

The mammalian cells that shall be provided for such a manufacturing may be epithelial cells in particular of the kidney, of the bowels system, of the liver or of the blood/brain barrier. Such cells can be also immortalized cells or recombinant cells. The cells can be taken from ordinary mammalian tissue or a primary cell culture.

Such double transfectant cells according to e] resulting from the manufacturing as mentioned above may be the cells of a cell culture as e.g. LLC-PK1 cells or MDCKII cells particularly when carrying recombinant vectors. Such LLC-PK1 cells may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hMRP2 protein (as exemplified by SEQ ID NO. 6). The hNTCP and the hMRP2 proteins may be expressed also simultaneously from the same vector construct. Such MDCKII may harbor for example a recombinant vector e.g. suitable for expressing the hNTCP protein (as exemplified by SEQ ID NO. 4) and a recombinant vector suitable for expressing the hMRP2 (as exemplified by SEQ ID NO. 6). Such a mammalian cell of a MDCKII type carrying two separate vectors with ability of expressing hNTCP protein and hMRP2 protein has been deposited with the DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1 b, D-38124 Braunschweig. The identification number is DSM ACC2644.

The invention pertains further to a monolayer of cells comprising at least two cells of mammalian cells having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hMRP2 protein.

Such a monolayer could occupy a part or the whole of a solid surface. The invention pertains to such a solid surface carrying a monolayer of such cells.
The solid surface could be formed by a plastic. The solid surface could also be part of a petri dish or a filter-insert.

The invention pertains also to a petri dish carrying a monolayer of cells as described and further to a filter-insert carrying a monolayer of cells as described. The membrane support of such a filter-insert could be made of polycarbonat and/or polyester. The pore size of the membrane support could be of about 0.4 µm +/- 0.2 µm. An according example of a filter-insert has been depicted in Fig. 1. There are several manufacturers of filter-inserts that could be used for placing the said monolayer of cells on it. Such manufacturers are for example “Corning Inc., Corning, NY” or “Millipore”.

The mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hMRP2 protein could be used e.g. for determining pharmacological profiles with respect to hepatobiliary elimination and/or renal excretion and/or brain resorption and/or intestinal resorption. For such use the mammalian cells could form part of a monolayer on a solid surface and/or a petri dish and/or on a filter-insert.

A protein shall be regarded as functional in context of this invention when it is in a condition to perform an activity in a biological context in particular as part of a living cell. Such an activity is detectable e.g. by an assy. A transporter is functional for example when this transporter moves a compound in particular the transporter's biological substrate from outside a cell into the inner compartment of this cell or vice versa. A biological substrate of an ion transporter protein consists e.g. of a monovalent and/or a divalent ion or other ions. The substrate of a glucose transporter protein is e.g. glucose. The substrate of a multiple drug resistance protein is e.g. a drug alone or conjugated to gluthathion or gluconate.

The handling of proteins in context of this invention can be achieved by a person skilled in the art by applying the according protocols from "Current Protocols in Protein Science" published by John Wiley & Sons (edited by: John E. Coligan, Ben M. Dunn, Hidde L. Ploegh, David W. Speicher, Paul T. Wingfield; 0-471-11184-8-Looseleaf; 0-471-14098-8-CDROM).

The handling of techniques concerning Molecular Biology as e.g. cloning, transforming of cells, sequencing, modifying promoters, expressing proteins or others can be achieved a the person skilled in the art by applying the according protocols from "Current Protocols in Molecular Biology" published by John Wiley & Sons (edited by: Fred M. Ausubel, Roger Brent, Robert E. Kingston, David D. Moore, J.G. Seidman, John A. Smith, Kevin Struhl; 0-471-50338-X-Looseleaf; 0-471-306614CDROM).

The handling of biological cells can be achieved a the person skilled in the art by applying the according protocols from "Current Protocols in Cell Biology" published by John Wiley & Sons (edited by: Juan S. Bonifacino, Mary Dasso, Jennifer Lippincott-Schwartz, Joe B. Harford, Kenneth M. Yamada; 0-471-24108-3-Looseleaf; 0-471-24105-9-CDROM).

Examples

Cell line generation: MDCKII-hBSEP and MDCKII-hNTCP/hBSEP:

MDCK II cells (Madin-Darby canine kidney cells, strain II) were cultured in Dulbecco's modified essential medium supplemented with 10 % fetal calf serum at 37°C, 5% CO_2, and 95% humidity. The MDCKII-hNTCP/hBSEP cell line (DSM ACC2643) was generated as follows: MDCK II cells were first transfected with the mammalian expression plasmid pcDNA3.1zeo(+)-hBSEP using the FuGene6 transfection reagent as described by the manufacturer (Roche). Transfected cells were selected with 1000µg/ml Zeocin. During selection the tissue culture medium was additionally supplemented with 100units/ml penicillin G and 100µg/ml streptomycin sulfate. The resulting cell clones were screened in the immunoblot for recombinant human BSEP expression followed by analysis of the proper subcellular localization of recombinant human BSEP using Confocal Laser Scanning Microscopy. The cell clone with the highest expression of human BSEP and concomitant localization of the recombinant BSEP at the apical membrane was chosen for FuGene6 transfection with the mammalian expression plasmid pcDNA3.1neo(+)-hNTCP. Transfected cells were selected with 800µg/ml Geneticin and 1000µg/ml Zeocin. During selection the tissue culture medium was additionally supplemented with 100units/ml penicillin G and 100µg/ml streptomycin sulfate. Double-transfected cell lines were again screened in the immunoblot for the expression of both, recombinant human BSEP and recombinant human NTCP, followed by analysis of the proper subcellular localization of both transgenes by Confocal Laser Scanning Microscopy. The cell clone with the highest expression and proper subcellular localization of both, human BSEP (apical membrane) and human NTCP (basolateral membrane) was chosen for further examinations.

Cell line generation: MDCKII-hMRP2 and MDCKII-hNTCP/hMRP2:

MDCK II cells (Madin-Darby canine kidney cells, strain II) were cultured in Dulbecco's modified essential medium supplemented with 10 % fetal calf serum at 37°C, 5% CO_2, and 95% humidity. The MDCKII-hNTCP/hMRP2 cell line (DSM ACC2644) was generated as follows: MDCK II cells were first transfected with the

mammalian expression plasmid pcDNA3.1zeo(+)-hMRP2 using the FuGene6 transfection reagent as described by the manufacturer (Roche). Transfected cells were selected with 1000µg/ml Zeocin. During selection the tissue culture medium was additionally supplemented with 100units/ml penicillin G and 100µg/ml streptomycin sulfate. The resulting cell clones were screened in the immunoblot for recombinant human MRP2 expression followed by analysis of the proper subcellular localization of recombinant human MRP2 using Confocal Laser Scanning Microscopy. The cell clone with the highest expression of human MRP2 and concomitant localization of the recombinant MRP2 at the apical membrane was chosen for FuGene6 transfection with the mammalian expression plasmid pcDNA3.1neo(+)-hNTCP. Transfected cells were selected with 800µg/ml Geneticin and 1000µg/ml Zeocin. During selection the tissue culture medium was additionally supplemented with 100units/ml penicillin G and 100µg/ml streptomycin sulfate. Double-transfected cell lines were again screened in the immunoblot for the expression of both, recombinant human MRP2 and recombinant human NTCP, followed by analysis of the proper subcellular localization of both transgenes by Confocal Laser Scanning Microscopy. The cell clone with the highest expression and proper subcellular localization of both, human MRP2 (apical membrane) and human NTCP (basolateral membrane) was chosen for further examinations.

Cell line generation: MDCKII-hNTCP:

MDCK II cells (Madin-Darby canine kidney cells, strain II) were cultured in Dulbecco's modified essential medium supplemented with 10 % fetal calf serum at 37°C, 5% CO_2, and 95% humidity. For generation of the MDCKII-hNTCP cell line MDCKII cells were transfected with the mammalian expression plasmid pcDNA3.1neo(+)-hNTCP using the FuGene6 transfection reagent as described by the manufacturer (Roche). Transfected cells were selected with 800µg/ml Geneticin. During selection the tissue culture medium was additionally supplemented with 100units/ml penicillin G and 100µg/ml streptomycin sulfate. The resulting cell clones were screened in the immunoblot for recombinant human NTCP expression followed by functional analysis of the cell clones with the highest human NTCP expression (measurement of taurocholic acid uptake in MDCKII-hNTCP cells compared with parental MDCKII cells).

Chemicals: [^{3}H(G)]-Bromosulfophthalein (BSP) (0.2 TBq/mmol) was purchased from Hartmann Analytic (Braunschweig, Germany). [2,4-^{3}H(N)]-Cholic acid (0.7 TBq/mmol), [Carboxyl-14C]-Chenodeoxycholic acid (1.9 GBq/mmol) and [Glycine-2-^{3}H]-Glycocholic acid (0.15 TBq/mmol) were obtained from Biotrend Chemikalien (Köln, Germany). [^{3}H(G)]-Taurocholic acid (0.1 TBq/mmol) was purchased from PerkinElmer Life Sciences. Zeocin, geneticin, penicillin G, and streptomycin sulfate were obtained from Invitrogen. Additional chemicals of analytical purity were obtained from Sigma and Merck (Darmstadt, Germany)

Transport assays:
Cells were grown at confluency (1x10^6 cells / insert) on polyester membrane inserts (pore size 0.4µM, insert diameter 24mm, Costar) in 6-well plates for 6 days. Overexpression of transgenes was induced by addition of 10mM sodium butyrate to the growth medium for 24 h. The next day cells were washed once with transport assay buffer (142mM NaCl, 5mM KCl, 5mM Glucose, 1.5mM $CaCl_2$, 1.2mM $MgSO_4$, 1mM KH_2PO_4, 12.5 mM HEPES, pH 7.3) followed by addition of [^{3}H]- or [^{14}C]-labeled compounds either to the apical (1.5 ml) or to the basolateral (2.5 ml) compartments. The opposite compartments were filled with transport assay buffer. After the indicated incubation periods (37°C) a sample was taken from every compartment and measured by scintillation counting (Wallac, Winspectral 1414 Liquid Scintillation Counter). The transcellular transport in percent of input was calculated as follows:

Transport from the apical to the basolateral compartment:

$$\text{Transport}_{A\text{-}B}\,[\%] = 100 \times (\text{Radioactivity}_{(t1)}\ \text{basolateral [dpm]} \times 2{,}5\ [\text{ml}])/(\text{Radioactivity}_{(t0)}\ \text{apical [dpm]} \times 1{,}5\ [\text{ml}])$$

Transport from the basolateral to the apical compartment:

$$\text{Transport}_{B\text{-}A}\,[\%] = 100 \times (\text{Radioactivity}_{(t1)}\ \text{apical [dpm]} \times 1{,}5\ [\text{ml}])/(\text{Radioactivity}_{(t0)}\ \text{basolateral [dpm]} \times 2{,}5\ [\text{ml}])$$

The following deposits of biological material were made in context of the present invention with DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b; D-38124 Braunschweig:

DSM ACC2643: Double transfected MDCKII cells harboring a vector for expressing human NTCP (Seq ID No. 4) and a vector for expressing human BSEP (Seq ID No. 5).

DSM ACC2644: Double transfected MDCKII cells harboring a vector for expressing human NTCP (Seq ID No. 4) and a vector for expressing human MRP2 (Seq ID No. 6).

Description of the Figures:

Fig 1:
General assay design
The cells are grown as a polarized and tight monolayers in filter inserts on a porous membrane (polyester membrane, pore size 0.4µM, 6-well) thereby separating a basolateral from an apical compartment. Fig. 1 is depicting an example of an filter-insert carrying a monolayer of cells.

Fig 2:
Assay design MDCK II-hNTCP/hBSEP
Schematic representation of MDCK II cells (A) and MDCK II cells transfected with human BSEP (B), or human NTCP (C), or both human NTCP and human BSEP (D). Human NTCP is localized at the basolateral membrane whereas human BSEP is localized at the apical membrane. The mono-transfected cell lines MDCKII-hBSEP and MDCKII-hNTCP serve as controls.

Fig 3:

Assay design MDCK II-hNTCP/hMRP2

Schematic representation of MDCK II cells (A) and MDCK II cells transfected with human MRP2 (B), or human NTCP (C), or both human NTCP and human MRP2 (D). Human NTCP is localized at the basolateral membrane whereas human MRP2 is localized at the apical membrane. The mono-transfected cell lines MDCKII-hMRP2 and MDCKII-hNTCP serve as controls.

Fig 4:

Vectorial transport of [^{3}H]Glycocholic acid, [^{3}H]Taurocholic acid, [^{3}H]Cholic acid, [^{14}C]Chenodeoxycholic acid and [^{3}H]BSP (Bromosulfophthalein) in MDCKII-hNTCP/hBSEP monolayers

MDCKII, MDCKII-hNTCP, MDCKII-hMRP2 and MDCKII-hNTCP/hMRP2 cells were grown on filter inserts for 6 days ($1x10^6$ cells/well). Transgene expression was induced by addition of 10mM sodium butyrate for 24 h. The indicated amount of each compound was given separately either to the basolateral or to the apical compartment. After 45 and 90 minutes an aliquot was taken from the opposite compartment and analysed by liquid scintillization counting. A-B indicates addition of compound to the apical compartment and sampling from the basolateral compartment (vectorial transport from the apical to the basolateral compartment). B-A indicates addition of compound to the basolateral compartment and sampling from the apical compartment (vectorial transport from the basolateral to the apical compartment). Data represent means ± SD (n=3).

Fig 5:

Vectorial transport of [^{3}H]BSP (Bromosulfophthalein), [^{3}H]Cholic acid and [^{3}H]Taurocholic acid in MDCKII-hNTCP/hMRP2 monolayers

MDCKII, MDCKII-hMRP2 and MDCKII-hNTCP/hMRP2 cells were grown on filter inserts for 6 days ($1x10^6$ cells/well). Transgene expression was induced by addition of 10mM sodium butyrate for 24 h. 100nM of each compound was given separately either to the basolateral or to the apical compartment. After one and two hours an aliquot was taken from the opposite compartment and analysed by liquid scintillization counting. A-B indicates addition of compound to the apical compartment

compartment and sampling from the basolateral compartment (vectorial transport from the apical to the basolateral compartment). B-A indicates addition of compound to the basolateral compartment and sampling from the apical compartment (vectorial transport from the basolateral to the apical compartment). Data represent means ± SD (n=3).

Fig 6:
Human NTCP (Slc10A1), cDNA sequence; Length: 1061 bases; the depicted sequence is corresponding to Seq ID No. 1.

Fig. 7:
Human BSEP (ABCB11), cDNA sequence; Length: 3966 bases; the depicted sequence is corresponding to Seq ID No. 2.

Fig. 8:
Human MRP2 (ABCC2), cDNA sequence; Length: 4650 bases; the depicted sequence is corresponding to Seq ID No. 3.

Fig. 9:
pcDNA3.1neo(+)-hNTCP, complete sequence; Length: 6433 bases; the depicted sequence is corresponding to Seq ID No. 4:
The human Sodium-dependent Tauchocholate Cotransporting Protein (hNTCP, SLC10A1) cDNA was cloned into the MCS of pcDNA3.1neo(+) (Invitrogen) via BamH I and Xba I. Human NTCP cDNA was PCR-amplified with gene-specific primers from a human liver cDNA library (Clontech). The cloned cDNA corresponds to GenBank Accession # L21893 from position 76-1136 (coding sequence: position 83-1132). Exceptions: position 616 c → t (no amino acid change) and position 774 t → c (amino acid change F → S).

Fig. 10:
pcDNA3.1zeo(+)-hBSEP, complete sequence; Length: 9043 bases; the depicted sequence is corresponding to Seq ID No. 5:

The human Bile Salt Export Pump (hBSEP, ABCB11) cDNA was cloned by using the Gateway Technology (Invitrogen). Human BSEP cDNA was PCR-amplified with attB-modified gene-specific primers from a human fetal liver cDNA library (Clontech) and cloned via pDONR221 into the pcDNA3.1zeo(+) vector (Invitrogen, modified to a Gateway destination vector). The cloned cDNA corresponds to GenBank Accession # AF136523 from position 128-4093 (coding sequence: position 128-4093). Exception: position 3211 G → A (no amino acid change).

Fig. 11:
pcDNA3.1zeo(+)-hMRP2, complete sequence; Length: 9658 bases; the depicted sequence is corresponding to Seq ID No. 6:
The human Multidrug-resistance Related Protein 2 (MRP2, ABCC2) cDNA was cloned into the MCS of pcDNA3.1zeo(+) (Invitrogen) via Nhe I and Afl II. Human MRP2 cDNA was PCR-amplified with gene-specific primers from a HepG2 cDNA library. The cloned cDNA corresponds to GenBank Accession # X96395 from position 26-4675 (coding sequence: position 38-4675). Exception: position 4009 C → T (no amino acid change).

Description of Seq IDs:

Seq ID No. 1:
Human NTCP (Slc10A1), cDNA sequence; Length: 1061 bases; Seq ID No. 1 is disclosed in Fig. 6.

Seq ID No. 2:
Human BSEP (ABCB11), cDNA sequence; Length: 3966 bases; Seq ID No. 2 is disclosed in Fig. 7.

Seq ID No. 3:
Human MRP2 (ABCC2), cDNA sequence; Length: 4650 bases; Seq ID No. 3 is disclosed in Fig. 8.

Seq ID No. 4:
pcDNA3.1neo(+)-hNTCP, complete sequence; Length: 6433 bases; Seq ID No. 4 is disclosed in Fig. 9.

Seq ID No. 5:
pcDNA3.1zeo(+)-hBSEP, complete sequence; Length: 9043 bases; Seq ID No. 5 is disclosed in Fig. 10.

Seq ID No. 6:
pcDNA3.1zeo(+)-hMRP2, complete sequence; Length: 9658 bases; Seq ID No. 6 is disclosed in Fig. 11.

References:

1. Louvard D. (1980) Apical membrane aminopeptidase appears at site of cell-cell contact in cultured kidney epithelial cells. Proc. Natl. Acad. Sci. USA 77(7): 4132-4136
2. Fuller S, von Bonsdorff C-H, and Simons K (1984) Vesicular Stomatitis Virus infects and matures only through the basolateral surface of the polarized epithelial cell line, MDCK. Cell 38: 65-77

Claims

1. Mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization at opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hBSEP protein.

2. Mammalian cell as claimed in claim 1 wherein the first side is the basolateral side and the second side is the apical side.

3. Mammalian cell as claimed in claim 1 and 2 wherein the first side is the apical side and the second side is the basolateral side.

4. Mammalian cell as claimed in claim 1 to 3 wherein the cell is an epithelial cell of the kidney, of the bowels system, of the liver or of the blood/brain barrier.

5. Mammalian cell as claimed in claims 1 to 4 which is immortalized.

6. Mammalian cell as claimed in claims 1 to 5 which is a recombinant cell.

7. Mammalian cell as claimed in claims 1 to 6 which is a LLC-PK1 cell harboring a vector for expressing hNTCP protein and a vector for expressing hBSEP protein.

8. Mammalian cell as claimed in claims 1 to 6 which is a MDCKII cell harboring a vector for expressing the hNTCP protein and a vector for expressing the hBSEP protein.

9. Mammalian cell as claimed in claim 8 according to deposit DSM ACC2643.

10. Manufacturing of a mammalian cell according to claims 1 to 9 wherein

a] a mammalian cell is provided;

b] a vector is provided encompassing the coding sequence of hNTCP;

c] a vector is provided encompassing the coding sequence of hBSEP;

d] the mammalian cell from a] is transformed by a vector from b] and by a vector from c] either simultaneously or consecutively;

e] a double transfectant cell from d] is identified and propagated.

11. Manufacturing of a mammalian cell as claimed in claim 10 wherein the mammalian cell from a] is an epithelial cell of the kidney, of the bowels system, of the liver or of the blood/brain barrier.

12. Manufacturing of a mammalian cell as claimed in claims 10 and 11 wherein the mammalian cell from a] is immortalized.

13. Manufacturing of a mammalian cell as claimed in claims 10 to 12 wherein the vector from b] is a polynucleotide according to Fig.9 (Seq ID No. 4).

14. Manufacturing of a mammalian cell as claimed in claims 10 to 12 wherein the vector from c] is a polynucleotide according to Fig. 10 (Seq ID No. 5).

15. Manufacturing of a mammalian cell as claimed in claims 10 to 14 wherein the mammalian cell is build up as deposited cell DSM ACC2643.

16. A monolayer of cells comprising at least two cells according to claims 1 to 9.

17. A solid surface carrying a monolayer according to claim 16 wherein the monolayer of cells could occupy the part or the whole of the solid surface.

18. A solid surface as claimed in claim 17 that is formed by a plastic.

19. A solid surface as claimed in claim 17 that is part of a petri dish.

20. A solid surface as claimed in claim 17 that is part of a filter-insert.

21. A petri dish carrying a monolayer of cells according to claim 17.

22. A filter-insert carrying a monolayer of cells according to claim 17.

23. A filter-insert as claimed in claim 22 wherein the membrane support is made of polycarbonat and/or polyester.

24. A filter-insert as claimed in claims 22 and 23 wherein the membrane support's pore size is 0.4 μm.

25. Use of a mammalian cell of claims 1 to 9 for determining a pharmacological profile with respect to hepatobiliary elimination and/or renal excretion and/or brain resorption and/or intestinal resorption.

26. Use of a mammalian cell as claimed in claim 25 wherein the mammalian cell forms part of a monolayer on a solid surface and/or on a petri dish and/or on a filter-insert.

27. Mammalian cell having a first and a second side which both sides form part of the outer surface of such cell and which both sides are different from the areas of contact of such cell and which first and second side are distinguished from each other by their localization of opposite ends of such cell wherein the first side carries a functional hNTCP protein and the second side carries a functional hMRP2 protein.

28. Mammalian cell as claimed in claim 27 wherein the first side is the basolateral side and the second side is the apical side.

29. Mammalian cell as claimed in claims 27 and 28 wherein the first side is the apical side and the second side is the basolateral side.

30. Mammalian cell as claimed in claims 27 to 29 wherein the cell is an epithelial cell of the kidney, of the bowels system, of the liver or of the blood/brain barrier.

31. Mammalian cell as claimed in claims 27 to 30 which is immortalized.

32. Mammalian cell as claimed in claims 27 to 31 which is a recombinant cell.

33. Mammalian cell as claimed in claims 27 to 32 which is a LLC-PK1 cell harboring a vector for expressing hNTCP protein and a vector for expressing hMRP2 protein.

34. Mammalian cell as claimed in claims 27 to 33 which is a MDCKII cell harboring a vector for expressing hNTCP protein and a vector for expressing hMRP2 protein.

35. Mammalian cell as claimed in claims 27 to 34 as deposited as DSM ACC2644.

36. Manufacturing of a mammalian cell according to claims 27 to 35 wherein

a] a mammalian cell is provided;
b] a vector is provided encompassing the coding sequence of hNTPC;
c] a vector is provided encompassing the coding sequence of hMRP2;
d] the mammalian cell from a] is transformed by a vector from b] and by a vector from c] either simultaneously or consecutively;
e] a double transfectant cell from d] is identified and propagated.

37. Manufacturing of a mammalian cell as claimed in claim 36 wherein the mammalian cell from a] is an epithelial cell of the kidney, of the bowels system, of the liver or of the blood/brain barrier.

38. Manufacturing of a mammalian cell as claimed in claims 36 and 37 wherein the mammalian cell from a] is immortalized.

39. Manufacturing of a mammalian cell as claimed in claims 36 to 38 wherein the vector from b] is a polynucleotide according to Fig. 9 (Seq ID No. 4).

40. Manufacturing of a mammalian cell as claimed in claims 36 to 38 wherein the the vector from c] is a polynucleotide according to Fig. 11 (Seq ID No. 6).

41. Manufacturing of a mammalian cell as claimed in claims 36 to 38 wherein the mammalian cell is build up as deposited cell DSM ACC2644.

42. A monolayer of cells comprising at least two cells according to claims 27 to 35.

43. A solid surface carrying a monolayer according to claim 42 wherein the monolayer of cells could occupy the part or the whole of the solid surface.

44. A solid surface as claimed in claim 43 that is formed by a plastic.

45. A solid surface as claimed in claim 43 that is part of a petri dish.

46. A solid surface as claimed in claim 43 that is part of a filter-insert.

47. A petri dish carrying a monolayer of cells according to claims 27 to 35.

48. A filter-insert carrying a monolayer of cells according to claims 27 to 35.

49. A filter-insert as claimed in claim 48 wherein the membrane support is made of polycarbonate and/or polyester.

50. A filter-insert as claimed in claim 48 wherein the membrane support's pore size is 0.4 μm.

51. Use of a mammalian cell of claims 27 to 35 for determining a pharmacological profile with respect to hepatobiliary elimination and/or renal excretion and/or brain resorption and/or intestinal resorption.

52. Use of a mammalian cell as claimed in claim 51 wherein the mammalian cell forms part of a monolayer on a solid surface and/or on a petri dish and/or on a filter-insert.

Summary

The invention is referring to several dual-transfected cell lines expressing human NTCP (Na/taurocholate Cotransporting Protein; SLC10A1) together with human BSEP (Bile Salt Export Pump; ABCB11) or human MRP2 (Multidrug Resistance Protein; ABCC2) suitable as an in vitro tool for pharmaceutical compound profiling particularly as a model for hepatobiliary elimination.

Fig. 1

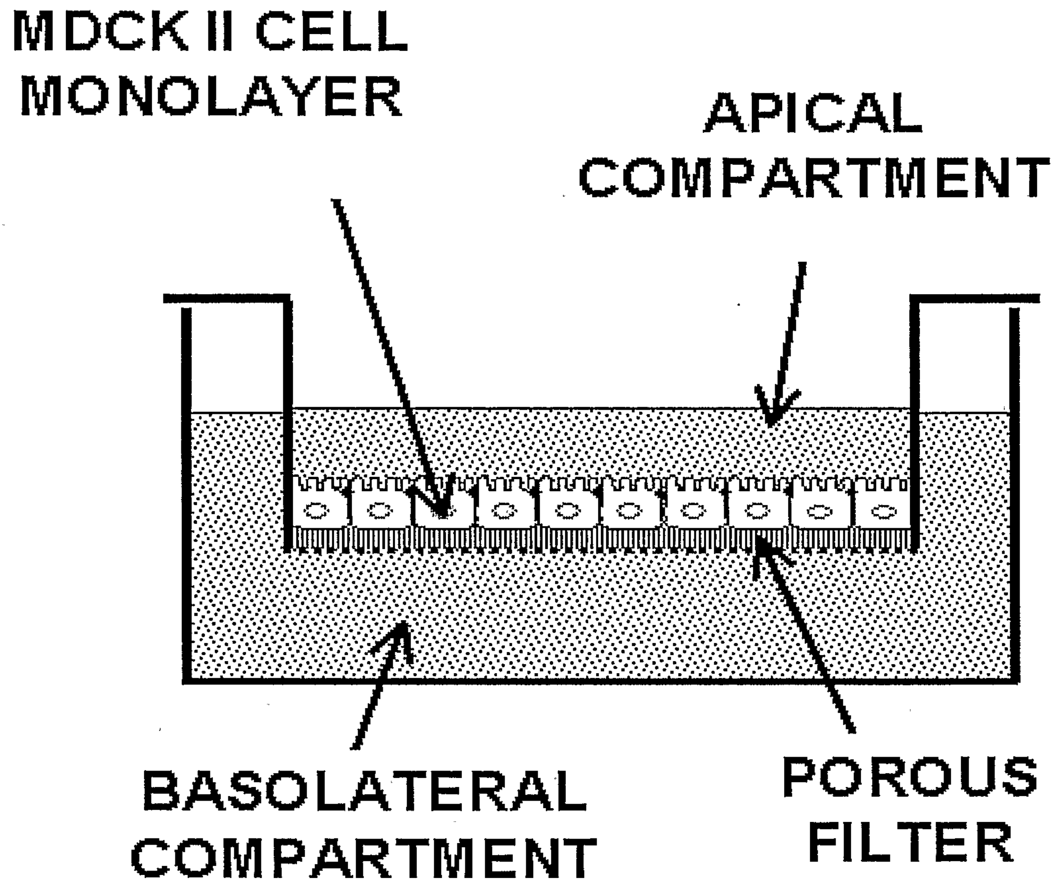

MDCK II CELL
MONOLAYER
APICAL
COMPARTMENT
BASOLATERAL
COMPARTMENT
POROUS
FILTER

Fig. 2

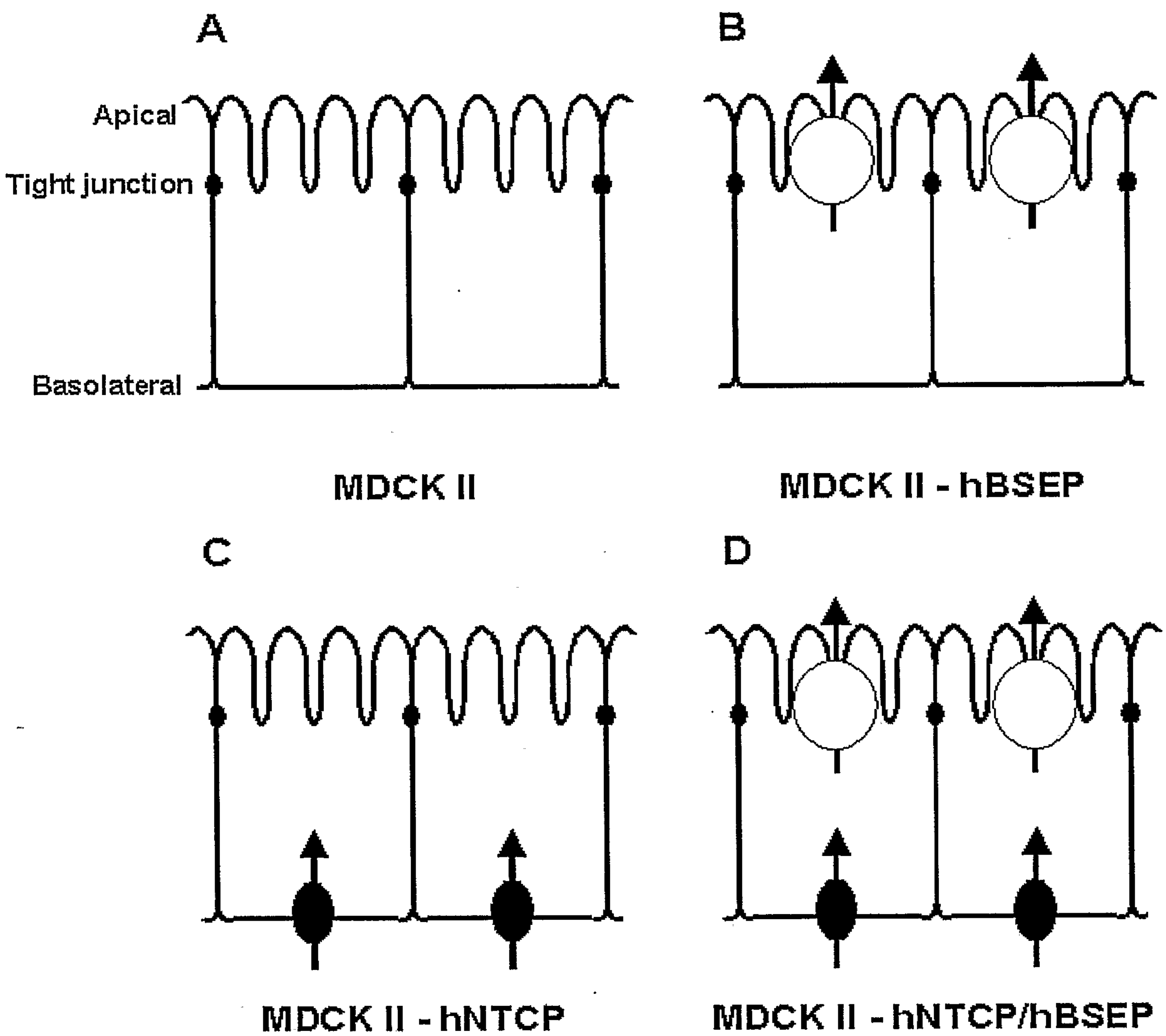

A
B
Apical
Tight junction
Basolateral
MDCK II
MDCK II - hBSEP
C
D
MDCK II - hNTCP
MDCK II - hNTCP/hBSEP

Fig 3

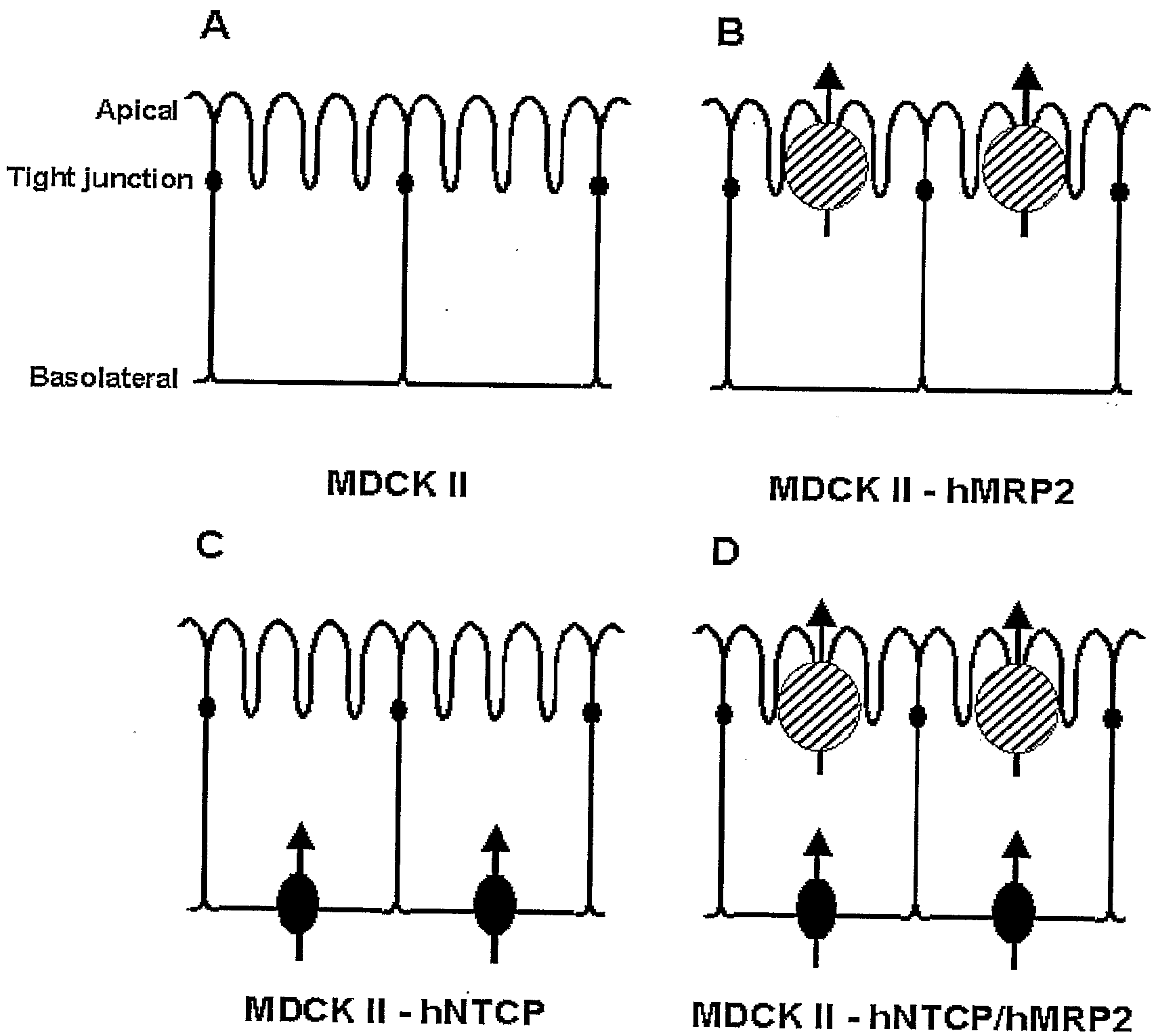
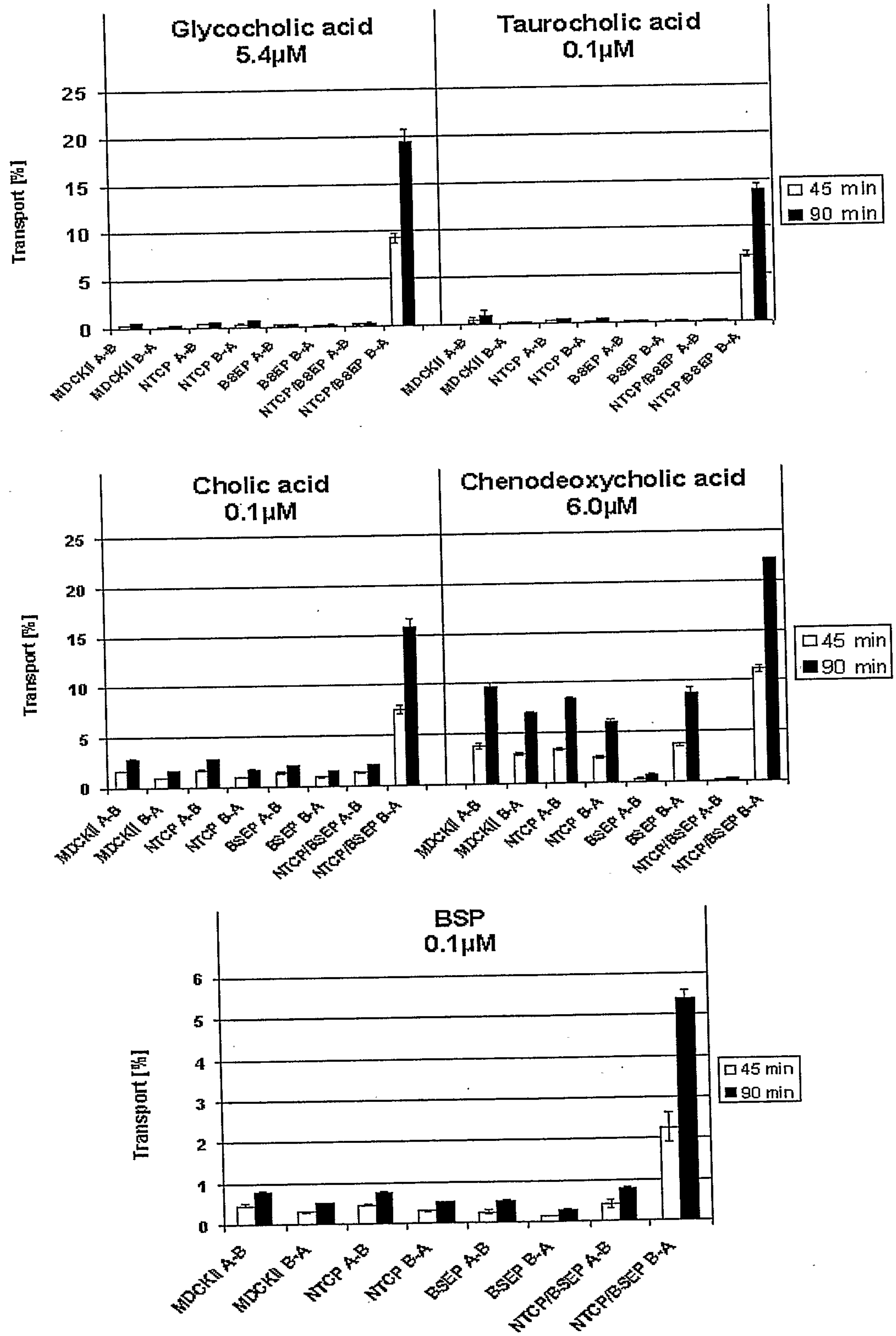

A
Apical
Tight junction
Basolateral
MDCK II
B
MDCK II - hMRP2
C
MDCK II - hNTCP
D
MDCK II - hNTCP/hMRP2

Fig 4



Glycocholic acid
5.4µM
Taurocholic acid
0.1µM
Transport [%]
0
5
10
15
20
25
45 min
90 min
MDCKII A-B
MDCKII B-A
NTCP A-B
NTCP B-A
BSEP A-B
BSEP B-A
NTCP/BSEP A-B
NTCP/BSEP B-A
Cholic acid
0.1µM
Chenodeoxycholic acid
6.0µM
BSP
0.1µM
0
1
2
3
4
5
6

Fig. 5

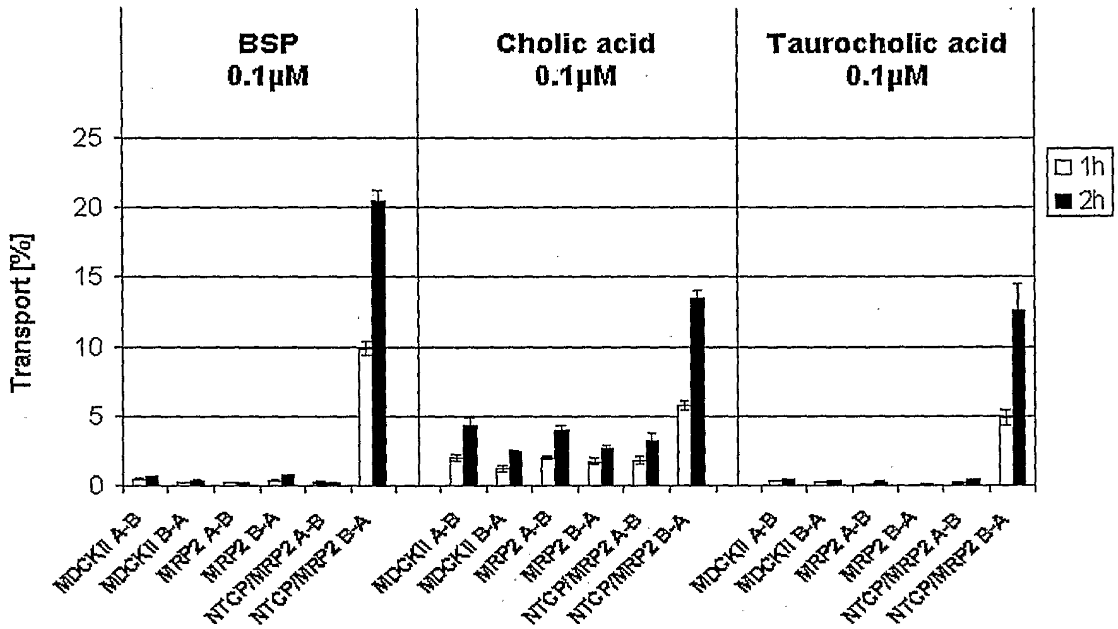

BSP
0.1µM
Cholic acid
0.1µM
Taurocholic acid
0.1µM
Transport [%]
25
20
15
10
5
0
□ 1h
■ 2h
MDCKII A-B
MDCKII B-A
MRP2 A-B
MRP2 B-A
NTCP/MRP2 A-B
NTCP/MRP2 B-A

Fig. 6

CAGGAGGATGGAGGCCCACAACGCGTCTGCCCCATTCAACTTCACCCTGCCACCCAACTTTGGCAAGCGCCCCACAGACC
TGGCACTGAGCGTCATCCTGGTGTTCATGTTGTTCTTCATCATGCTCTCGCTGGGCTGCACCATGGAGTTCAGCAAGATC
AAGGCTCACTTATGGAAGCCTAAAGGGCTGGCCATCGCCCTGGTGGCACAGTATGGCATCATGCCCCTCACGGCCTTTGT
GCTGGGCAAGGTCTTCCGGCTGAAGAACATTGAGGCACTGGCCATCTTGGTCTGTGGCTGCTCACCTGGAGGGAACCTGT
CCAATGTCTTCAGTCTGGCCATGAAGGGGGACATGAACCTCAGCATTGTGATGACCACCTGCTCCACCTTCTGTGCCCTT
GGCATGATGCCTCTCCTCCTGTACATCTACTCCAGGGGGATCTATGATGGGGACCTGAAGGACAAGGTGCCCTATAAAGG
CATCGTGATATCACTGGTCCTGGTTCTCATTCCTTGCACCATAGGGATCGTCCTCAAATCTAAACGGCCACAATACATGC
GCTATGTCATCAAGGGAGGGATGATCATCATTCTCTTGTGCAGTGTGGCCGTCACAGTTCTCTCTGCCATCAATGTGGGG
AAGAGCATCATGTTTGCCATGACACCACTCTTGATTGCCACCTCCTCCCTGATGCCTTCTATTGGCTTTCTGCTGGGTTA
TGTTCTCTCTGCTCTCTTCTGCCTCAATGGACGGTGCAGACGCACTGTCAGCATGGAGACTGGATGCCAAAATGTCCAAC
TCTGTTCCACCATCCTCAATGTGGCCTTTCCACCTGAAGTCATTGGACCACTTTTCTTCTTTCCCCTCCTCTACATGATT
TTCCAGCTTGGAGAAGGGCTTCTCCTCATTGCCATATTTTGGTGCTATGAGAAATTCAAGACTCCCAAGGATAAAACAAA
AATGATCTACACAGCTGCCACAACTGAAGAAACAATTCCAGGAGCTCTGGGAAATGGCACCTACAAAGGGGAGGACTGCT
CCCCTTGCACAGCCTAGCCCT

Fig. 7

ATGTCTGACTCAGTAATTCTTCGAAGTATAAAGAAATTTGGAGAGGAGAATGATGGTTTTGAGTCAGATAAATCATATAA
TAATGATAAGAAATCAAGGTTACAAGATGAGAAGAAAGGTGATGGCGTTAGAGTTGGCTTCTTTCAATTGTTTCGGTTTT
CTTCATCAACTGACATTTGGCTGATGTTTGTGGGAAGTTTGTGTGCATTTCTCCATGGAATAGCCCAGCCAGGCGTGCTA
CTCATTTTTGGCACAATGACAGATGTTTTTATTGACTACGACGTTGAGTTACAAGAACTCCAGATTCCAGGAAAAGCATG
TGTGAATAACACCATTGTATGGACTAACAGTTCCCTCAACCAGAACATGACAAATGGAACACGTTGTGGGTTGCTGAACA
TCGAGAGCGAAATGATCAAATTTGCCAGTTACTATGCTGGAATTGCTGTCGCAGTACTTATCACAGGATATATTCAAATA
TGCTTTTGGGTCATTGCCGCAGCTCGTCAGATACAGAAAATGAGAAAATTTTACTTTAGGAGAATAATGAGAATGGAAAT
AGGGTGGTTTGACTGCAATTCAGTGGGGGAGCTGAATACAAGATTCTCTGATGATATTAATAAAATCAATGATGCCATAG
CTGACCAAATGGCCCTTTTCATTCAGCGCATGACCTCGACCATCTGTGGTTTCCTGTTGGGATTTTTCAGGGGTTGGAAA
CTGACCTTGGTTATTATTTCTGTCAGCCCTCTCATTGGGATTGGAGCAGCCACCATTGGTCTGAGTGTGTCCAAGTTTAC
GGACTATGAGCTGAAGGCCTATGCCAAAGCAGGGGTGGTGGCTGATGAAGTCATTTCATCAATGAGAACAGTGGCTGCTT
TTGGTGGTGAGAAAAGAGAGGTTGAAAGGTATGAGAAAAATCTTGTGTTCGCCCAGCGTTGGGGAATTAGAAAAGGAATA
GTGATGGGATTCTTTACTGGATTCGTGTGGTGTCTCATCTTTTTGTGTTATGCACTGGCCTTCTGGTACGGCTCCACACT
TGTCCTGGATGAAGGAGAATATACACCAGGAACCCTTGTCCAGATTTTCCTCAGTGTCATAGTAGGAGCTTTAAATCTTG
GCAATGCCTCTCCTTGTTTGGAAGCCTTTGCAACTGGACGTGCAGCAGCCACCAGCATTTTTGAGACAATAGACAGGAAA
CCCATCATTGACTGCATGTCAGAAGATGGTTACAAGTTGGATCGAATCAAGGGTGAAATTGAATTCCATAATGTGACCTT
CCATTATCCTTCCAGACCAGAGGTGAAGATTCTAAATGACCTCAACATGGCCATTAAACCAGGGGAAATGACAGCTCTGG
TAGGACCCAGTGGAGCTGGAAAAAGTACAGCACTGCAACTCATTCAGCGATTCTATGACCCCTGTGAAGGAATGGTGACC
GTGGATGGCCATGACATTCGCTCTCTTAACATTCAGTGGCTTAGAGATCAGATTGGGATAGTGGAGCAAGAGCCAGTTCT
GTTCTCTACCACCATTGCAGAAAATATTCGCTATGGCAGAGAAGATGCAACAATGGAAGACATAGTCCAAGCTGCCAAGG
AGGCCAATGCCTACAACTTCATCATGGACCTGCCACAGCAATTTGACACCCTTGTTGGAGAAGGAGGAGGCCAGATGAGT
GGTGGCCAGAAACAAAGGGTAGCTATCGCCAGAGCCCTCATCCGAAATCCCAAGATTCTGCTTTTGGACATGGCCACCTC
AGCTCTGGACAATGAGAGTGAAGCCATGGTGCAAGAAGTGCTGAGTAAGATTCAGCATGGGCACACAATCATTTCAGTTG

CTCATCGCTTGTCTACGGTCAGAGCTGCAGATACCATCATTGGTTTTGAACATGGCACTGCAGTGGAAAGAGGGACCCAT
GAAGAATTACTGGAAAGGAAAGGTGTTTACTTCACTCTAGTGACTTTGCAAAGCCAGGGAAATCAAGCTCTTAATGAAGA
GGACATAAAGGATGCAACTGAAGATGACATGCTTGCGAGGACCTTTAGCAGAGGGAGCTACCAGGATAGTTTAAGGGCTT
CCATCCGGCAACGCTCCAAGTCTCAGCTTTCTTACCTGGTGCACGAACCTCCATTAGCTGTTGTAGATCATAAGTCTACC
TATGAAGAAGATAGAAAGGACAAGGACATTCCTGTGCAGGAAGAAGTTGAACCTGCCCCAGTTAGGAGGATTCTGAAATT
CAGTGCTCCAGAATGGCCCTACATGCTGGTAGGGTCTGTGGGTGCAGCTGTGAACGGGACAGTCACACCCTTGTATGCCT
TTTTATTCAGCCAGATTCTTGGGACTTTTTCAATTCCTGATAAAGAGGAACAAAGGTCACAGATCAATGGTGTGTGCCTA
CTTTTTGTAGCAATGGGCTGTGTATCTCTTTTCACCCAATTTCTACAGGGATATGCCTTTGCTAAATCTGGGGAGCTCCT
AACAAAAAGGCTACGTAAATTTGGTTTCAGGGCAATGCTGGGGCAAGATATTGCCTGGTTTGATGACCTCAGAAATAGCC
CTGGAGCATTGACAACAAGACTTGCTACAGATGCTTCCCAAGTTCAAGGGGCTGCCGGCTCTCAGATCGGGATGATAGTC
AATTCCTTCACTAACGTCACTGTGGCCATGATCATTGCCTTCTCCTTTAGCTGGAAGCTGAGCCTGGTCATCTTGTGCTT
CTTCCCCTTCTTGGCTTTATCAGGAGCCACACAGACCAGGATGTTGACAGGATTTGCCTCTCGAGATAAGCAGGCCCTGG
AGATGGTGGGACAGATTACAAATGAAGCCCTCAGTAACATCCGCACTGTTGCTGGAATTGGAAAGGAGAGGCGGTTCATT
GAAGCACTTGAGACTGAGCTGGAGAAGCCCTTCAAGACAGCCATTCAGAAAGCCAATATTTACGGATTCTGCTTTGCCTT
TGCCCAGTGCATCATGTTTATTGCGAATTCTGCTTCCTACAGATATGGAGGTTACTTAATCTCCAATGAGGGGCTCCATT
TCAGCTATGTGTTCAGGGTGATCTCTGCAGTTGTACTGAGTGCAACAGCTCTTGGAAGAGCCTTCTCTTACACCCCAAGT
TATGCAAAAGCTAAAATATCAGCTGCACGCTTTTTCAACTGCTGGACCGACAACCCCCAATCAGTGTATACAATACTGC
AGGTGAAAAATGGGACAACTTCCAGGGGAAGATTGATTTTGTTGATTGTAAATTTACATATCCTTCTCGACCTGACTCGC
AAGTTCTGAATGGTCTCTCAGTGTCGATTAGTCCAGGGCAGACACTGGCGTTTGTTGGGAGCAGTGGATGTGGCAAAAGC
ACTAGCATTCAGCTGTTGGAACGTTTCTATGATCCTGATCAAGGGAAGGTGATGATAGATGGTCATGACAGCAAAAAAGT
AAATGTCCAGTTCCTCCGCTCAAACATTGGAATTGTTTCCCAGGAACCAGTGTTGTTTGCCTGTAGCATAATGGACAATA
TCAAGTATGGAGACAACACCAAAGAAATTCCCATGGAAAGAGTCATAGCAGCTGCAAAACAGGCTCAGCTGCATGATTTT
GTCATGTCACTCCCAGAGAAATATGAAACTAACGTTGGGTCCCAGGGGTCTCAACTCTCTAGAGGGGAGAAACAACGCAT
TGCTATTGCTCGGGCCATTGTACGAGATCCTAAAATCTTGCTACTAGATGAAGCCACTTCTGCCTTAGACACAGAAAGTG
AAAAGACGGTGCAGGTTGCTCTAGACAAAGCCAGAGAGGGTCGGACCTGCATTGTCATTGCCCATCGCTTGTCCACCATC
CAGAACGCGGATATCATTGCTGTCATGGCACAGGGGGTGGTGATTGAAAAGGGGACCCATGAAGAACTGATGGCCCAAAA
AGGAGCCTACTACAAACTAGTCACCACTGGATCCCCCATCAGTTGA

Fig. 8

AGTCCAGGAATCATGCTGGAGAAGTTCTGCAACTCTACTTTTTGGAATTCCTCATTCCTGGACAGTCCGGAGGCAGACCT
GCCACTTTGTTTTGAGCAAACTGTTCTGGTGTGGATTCCCTTGGGCTTCCTATGGCTCCTGGCCCCCTGGCAGCTTCTCC
ACGTGTATAAATCCAGGACCAAGAGATCCTCTACCACCAAACTCTATCTTGCTAAGCAGGTATTCGTTGGTTTTCTTCTT
ATTCTAGCAGCCATAGAGCTGGCCCTTGTACTCACAGAAGACTCTGGACAAGCCACAGTCCCTGCTGTTCGATATACCAA
TCCAAGCCTCTACCTAGGCACATGGCTCCTGGTTTTGCTGATCCAATACAGCAGACAATGGTGTGTACAGAAAAACTCCT
GGTTCCTGTCCCTATTCTGGATTCTCTCGATACTCTGTGGCACTTTCCAATTTCAGACTCTGATCCGGACACTCTTACAG
GGTGACAATTCTAATCTAGCCTACTCCTGCCTGTTCTTCATCTCCTACGGATTCCAGATCCTGATCCTGATCTTTTCAGC
ATTTTCAGAAAATAATGAGTCATCAAATAATCCATCATCCATAGCTTCATTCCTGAGTAGCATTACCTACAGCTGGTATG
ACAGCATCATTCTGAAAGGCTACAAGCGTCCTCTGACACTCGAGGATGTCTGGGAAGTTGATGAAGAGATGAAAACCAAG
ACATTAGTGAGCAAGTTTGAAACGCACATGAAGAGAGAGCTGCAGAAAGCCAGGCGGGCACTCCAGAGACGGCAGGAGAA
GAGCTCCCAGCAGAACTCTGGAGCCAGGCTGCCTGGCTTGAACAAGAATCAGAGTCAAAGCCAAGATGCCCTTGTCCTGG
AAGATGTTGAAAAGAAAAAAAAGAAGTCTGGGACCAAAAAAGATGTTCCAAAATCCTGGTTGATGAAGGCTCTGTTCAAA
ACTTTCTACATGGTGCTCCTGAAATCATTCCTACTGAAGCTAGTGAATGACATCTTCACGTTTGTGAGTCCTCAGCTGCT
GAAATTGCTGATCTCCTTTGCAAGTGACCGTGACACATATTTGTGGATTGGATATCTCTGTGCAATCCTCTTATTCACTG

CGGCTCTCATTCAGTCTTTCTGCCTTCAGTGTTATTTCCAACTGTGCTTCAAGCTGGGTGTAAAAGTACGGACAGCTATC
ATGGCTTCTGTATATAAGAAGGCATTGACCCTATCCAACTTGGCCAGGAAGGAGTACACCGTTGGAGAAACAGTGAACCT
GATGTCTGTGGATGCCCAGAAGCTCATGGATGTGACCAACTTCATGCACATGCTGTGGTCAAGTGTTCTACAGATTGTCT
TATCTATCTTCTTCCTATGGAGAGAGTTGGGACCCTCAGTCTTAGCAGGTGTTGGGGTGATGGTGCTTGTAATCCCAATT
AATGCGATACTGTCCACCAAGAGTAAGACCATTCAGGTCAAAAATATGAAGAATAAAGACAAACGTTTAAAGATCATGAA
TGAGATTCTTAGTGGAATCAAGATCCTGAAATATTTTGCCTGGGAACCTTCATTCAGAGACCAAGTACAAAACCTCCGGA
AGAAAGAGCTCAAGAACCTGCTGGCCTTTAGTCAACTACAGTGTGTAGTAATATTCGTCTTCCAGTTAACTCCAGTCCTG
GTATCTGTGGTCACATTTTCTGTTTATGTCCTGGTGGATAGCAACAATATTTTGGATGCACAAAAGGCCTTCACCTCCAT
TACCCTCTTCAATATCCTGCGCTTTCCCCTGAGCATGCTTCCCATGATGATCTCCTCCATGCTCCAGGCCAGTGTTTCCA
CAGAGCGGCTAGAGAAGTACTTGGGAGGGGATGACTTGGACACATCTGCCATTCGACATGACTGCAATTTTGACAAAGCC
ATGCAGTTTTCTGAGGCCTCCTTTACCTGGGAACATGATTCGGAAGCCACAGTCCGAGATGTGAACCTGGACATTATGGC
AGGCCAACTTGTGGCTGTGATAGGCCCTGTCGGCTCTGGGAAATCCTCCTTGATATCAGCCATGCTGGGAGAAATGGAAA
ATGTCCACGGGCACATCACCATCAAGGGCACCACTGCCTATGTCCCACAGCAGTCCTGGATTCAGAATGGCACCATAAAG
GACAACATCCTTTTTGGAACAGAGTTTAATGAAAAGAGGTACCAGCAAGTACTGGAGGCCTGTGCTCTCCTCCCAGACTT
GGAAATGCTGCCTGGAGGAGATTTGGCTGAGATTGGAGAGAAGGGTATAAATCTTAGTGGGGGTCAGAAGCAGCGGATCA
GCCTGGCCAGAGCTACCTACCAAAATTTAGACATCTATCTTCTAGATGACCCCCTGTCTGCAGTGGATGCTCATGTAGGA
AAACATATTTTTAATAAGGTCTTGGGCCCCAATGGCCTGTTGAAAGGCAAGACTCGACTCTTGGTTACACATAGCATGCA
CTTTCTTCCTCAAGTGGATGAGATTGTAGTTCTGGGGAATGGAACAATTGTAGAGAAAGGATCCTACAGTGCTCTCCTGG
CCAAAAAAGGAGAGTTTGCTAAGAATCTGAAGACATTTCTAAGACATACAGGCCCTGAAGAGGAAGCCACAGTCCATGAT
GGCAGTGAAGAAGAAGACGATGACTATGGGCTGATATCCAGTGTGGAAGAGATCCCCGAAGATGCAGCCTCCATAACCAT
GAGAAGAGAGAACAGCTTTCGTCGAACACTTAGCCGCAGTTCTAGGTCCAATGGCAGGCATCTGAAGTCCCTGAGAAACT
CCTTGAAAACTCGGAATGTGAATAGCCTGAAGGAAGACGAAGAACTAGTGAAAGGACAAAAACTAATTAAGAAGGAATTC
ATAGAAACTGGAAAGGTGAAGTTCTCCATCTACCTGGAGTACCTACAAGCAATAGGATTGTTTTCGATATTCTTCATCAT
CCTTGCGTTTGTGATGAATTCTGTGGCTTTTATTGGATCCAACCTCTGGCTCAGTGCTTGGACCAGTGACTCTAAAATCT
TCAATAGCACCGACTATCCAGCATCTCAGAGGGACATGAGAGTTGGAGTCTACGGAGCTCTGGGATTAGCCCAAGGTATA
TTTGTGTTCATAGCACATTTCTGGAGTGCCTTTGGTTTCGTCCATGCATCAAATATCTTGCACAAGCAACTGCTGAACAA
TATCCTTCGAGCACCTATGAGATTTTTTGACACAACACCCACAGGCCGGATTGTGAACAGGTTTGCCGGCGATATTTCCA
CAGTGGATGACACCCTGCCTCAGTCCTTGCGCAGCTGGATTACATGCTTCCTGGGGATAATCAGCACCCTTGTCATGATC
TGCATGGCCACTCCTGTCTTCACCATCATCGTCATTCCTCTTGGCATTATTTATGTATCTGTTCAGATGTTTTATGTGTC
TACCTCCCGCCAGCTGAGGCGTCTGGACTCTGTCACCAGGTCCCCAATCTACTCTCACTTCAGCGAGACCGTATCAGGTT
TGCCAGTTATCCGTGCCTTTGAGCACCAGCAGCGATTTCTGAAACACAATGAGGTGAGGATTGACACCAACCAGAAATGT
GTCTTTTCCTGGATCACCTCCAACAGGTGGCTTGCAATTCGCCTGGAGCTGGTTGGGAACCTGACTGTCTTCTTTTCAGC
CTTGATGATGGTTATTTATAGAGATACCCTAAGTGGGGACACTGTTGGCTTTGTTCTGTCCAATGCACTCAATATCACAC
AAACCCTGAACTGGCTGGTGAGGATGACATCAGAAATAGAGACCAACATTGTGGCTGTTGAGCGAATAACTGAGTACACA
AAAGTGGAAAATGAGGCACCCTGGGTGACTGATAAGAGGCCTCCGCCAGATTGGCCCAGCAAAGGCAAGATCCAGTTTAA
CAACTACCAAGTGCGGTACCGACCTGAGCTGGATCTGGTCCTCAGAGGGATCACTTGTGACATTGGTAGCATGGAGAAGA
TTGGTGTGGTGGGCAGGACAGGAGCTGGAAAGTCATCCCTCACAAACTGCCTCTTCAGAATCTTAGAGGCTGCCGGTGGT
CAGATTATCATTGATGGAGTAGATATTGCTTCCATTGGGCTCCACGACCTCCGAGAGAAGCTGACCATCATCCCCCAGGA
CCCCATCCTGTTCTCTGGAAGCCTGAGGATGAATCTCGACCCTTTCAACAACTACTCAGATGAGGAGATTTGGAAGGCCT
TGGAGCTGGCTCACCTCAAGTCTTTTGTGGCCAGCCTGCAACTTGGGTTATCCCACGAAGTGACAGAGGCTGGTGGCAAC
CTGAGCATAGGCCAGAGGCAGCTGCTGTGCCTGGGCAGGGCTCTGCTTCGGAAATCCAAGATCCTGGTCCTGGATGAGGC
CACTGCTGCGGTGGATCTAGAGACAGACAACCTCATTCAGACGACCATCCAAAACGAGTTCGCCCACTGCACAGTGATCA
CCATCGCCCACAGGCTGCACACCATCATGGACAGTGACAAGGTAATGGTCCTAGACAACGGGAAGATTATAGAGTGCGGC
AGCCCTGAAGAACTGCTACAAATCCCTGGACCCTTTTACTTTATGGCTAAGGAAGCTGGCATTGAGAATGTGAACAGCAC
AAAATTCTAG

Fig. 9

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTAT
CTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATT
GATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAA
CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
TGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAA
TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCCAGGAGGATGGAGGCCCACAACGCGT
CTGCCCCATTCAACTTCACCCTGCCACCCAACTTTGGCAAGCGCCCCACAGACCTGGCACTGAGCGTCATCCTGGTGTTC
ATGTTGTTCTTCATCATGCTCTCGCTGGGCTGCACCATGGAGTTCAGCAAGATCAAGGCTCACTTATGGAAGCCTAAAGG
GCTGGCCATCGCCCTGGTGGCACAGTATGGCATCATGCCCCTCACGGCCTTTGTGCTGGGCAAGGTCTTCCGGCTGAAGA
ACATTGAGGCACTGGCCATCTTGGTCTGTGGCTGCTCACCTGGAGGGAACCTGTCCAATGTCTTCAGTCTGGCCATGAAG
GGGGACATGAACCTCAGCATTGTGATGACCACCTGCTCCACCTTCTGTGCCCTTGGCATGATGCCTCTCCTCCTGTACAT
CTACTCCAGGGGGATCTATGATGGGGACCTGAAGGACAAGGTGCCCTATAAAGGCATCGTGATATCACTGGTCCTGGTTC
TCATTCCTTGCACCATAGGGATCGTCCTCAAATCTAAACGGCCACAATACATGCGCTATGTCATCAAGGGAGGGATGATC
ATCATTCTCTTGTGCAGTGTGGCCGTCACAGTTCTCTCTGCCATCAATGTGGGGAAGAGCATCATGTTTGCCATGACACC
ACTCTTGATTGCCACCTCCTCCCTGATGCCTTCTATTGGCTTTCTGCTGGGTTATGTTCTCTCTGCTCTCTTCTGCCTCA
ATGGACGGTGCAGACGCACTGTCAGCATGGAGACTGGATGCCAAAATGTCCAACTCTGTTCCACCATCCTCAATGTGGCC
TTTCCACCTGAAGTCATTGGACCACTTTTCTTCTTTCCCCTCCTCTACATGATTTTCCAGCTTGGAGAAGGGCTTCTCCT
CATTGCCATATTTTGGTGCTATGAGAAATTCAAGACTCCCAAGGATAAAACAAAAATGATCTACACAGCTGCCACAACTG
AAGAAACAATTCCAGGAGCTCTGGGAAATGGCACCTACAAAGGGGAGGACTGCTCCCCTTGCACAGCCTAGCCCTTCTAG
AGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC
CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGG
TGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCG
CATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCT
TTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGG
TTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATC
TCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATT
TAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCA
AAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTG
AGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGA
CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTC
GGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCT

TTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGG
GCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAG
GATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGA
TCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCG
ATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGAC
GGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCAT
CGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCG
GCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT
GACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGAT
TCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGA
TCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAA
ATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGT
ATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAAT
TCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
TGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC
GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC
AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAAT
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG
CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT
CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC
CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG
TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA
GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA
ACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
TCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAC
CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG
ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCA
TGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATG
GTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA
GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGT
TCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATT
GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCG
CGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Fig. 10

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTAT
CTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATT
GATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAA
CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
TGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAA
TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTATCACAAGTTTGTACAAAAAAGCAGGCTTAGGAATGTCTGACTCA
GTAATTCTTCGAAGTATAAAGAAATTTGGAGAGGAGAATGATGGTTTTGAGTCAGATAAATCATATAATAATGATAAGAA
ATCAAGGTTACAAGATGAGAAGAAAGGTGATGGCGTTAGAGTTGGCTTCTTTCAATTGTTTCGGTTTTCTTCATCAACTG
ACATTTGGCTGATGTTTGTGGGAAGTTTGTGTGCATTTCTCCATGGAATAGCCCAGCCAGGCGTGCTACTCATTTTTGGC
ACAATGACAGATGTTTTTATTGACTACGACGTTGAGTTACAAGAACTCCAGATTCCAGGAAAAGCATGTGTGAATAACAC
CATTGTATGGACTAACAGTTCCCTCAACCAGAACATGACAAATGGAACACGTTGTGGGTTGCTGAACATCGAGAGCGAAA
TGATCAAATTTGCCAGTTACTATGCTGGAATTGCTGTCGCAGTACTTATCACAGGATATATTCAAATATGCTTTTGGGTC
ATTGCCGCAGCTCGTCAGATACAGAAAATGAGAAAATTTTACTTTAGGAGAATAATGAGAATGGAAATAGGGTGGTTTGA
CTGCAATTCAGTGGGGGAGCTGAATACAAGATTCTCTGATGATATTAATAAAATCAATGATGCCATAGCTGACCAAATGG
CCCTTTTCATTCAGCGCATGACCTCGACCATCTGTGGTTTCCTGTTGGGATTTTTCAGGGGTTGGAAACTGACCTTGGTT
ATTATTTCTGTCAGCCCTCTCATTGGGATTGGAGCAGCCACCATTGGTCTGAGTGTGTCCAAGTTTACGGACTATGAGCT
GAAGGCCTATGCCAAAGCAGGGGTGGTGGCTGATGAAGTCATTTCATCAATGAGAACAGTGGCTGCTTTTGGTGGTGAGA
AAAGAGAGGTTGAAAGGTATGAGAAAAATCTTGTGTTCGCCCAGCGTTGGGGAATTAGAAAAGGAATAGTGATGGGATTC
TTTACTGGATTCGTGTGGTGTCTCATCTTTTTGTGTTATGCACTGGCCTTCTGGTACGGCTCCACACTTGTCCTGGATGA
AGGAGAATATACACCAGGAACCCTTGTCCAGATTTTCCTCAGTGTCATAGTAGGAGCTTTAAATCTTGGCAATGCCTCTC
CTTGTTTGGAAGCCTTTGCAACTGGACGTGCAGCAGCCACCAGCATTTTTGAGACAATAGACAGGAAACCCATCATTGAC
TGCATGTCAGAAGATGGTTACAAGTTGGATCGAATCAAGGGTGAAATTGAATTCCATAATGTGACCTTCCATTATCCTTC
CAGACCAGAGGTGAAGATTCTAAATGACCTCAACATGGCCATTAAACCAGGGGAAATGACAGCTCTGGTAGGACCCAGTG
GAGCTGGAAAAAGTACAGCACTGCAACTCATTCAGCGATTCTATGACCCCTGTGAAGGAATGGTGACCGTGGATGGCCAT
GACATTCGCTCTCTTAACATTCAGTGGCTTAGAGATCAGATTGGGATAGTGGAGCAAGAGCCAGTTCTGTTCTCTACCAC
CATTGCAGAAAATATTCGCTATGGCAGAGAAGATGCAACAATGGAAGACATAGTCCAAGCTGCCAAGGAGGCCAATGCCT
ACAACTTCATCATGGACCTGCCACAGCAATTTGACACCCTTGTTGGAGAAGGAGGAGGCCAGATGAGTGGTGGCCAGAAA
CAAAGGGTAGCTATCGCCAGAGCCCTCATCCGAAATCCCAAGATTCTGCTTTTGGACATGGCCACCTCAGCTCTGGACAA
TGAGAGTGAAGCCATGGTGCAAGAAGTGCTGAGTAAGATTCAGCATGGGCACACAATCATTTCAGTTGCTCATCGCTTGT
CTACGGTCAGAGCTGCAGATACCATCATTGGTTTTGAACATGGCACTGCAGTGGAAAGAGGGACCCATGAAGAATTACTG
GAAAGGAAAGGTGTTTACTTCACTCTAGTGACTTTGCAAAGCCAGGGAAATCAAGCTCTTAATGAAGAGGACATAAAGGA
TGCAACTGAAGATGACATGCTTGCGAGGACCTTTAGCAGAGGGAGCTACCAGGATAGTTTAAGGGCTTCCATCCGGCAAC
GCTCCAAGTCTCAGCTTTCTTACCTGGTGCACGAACCTCCATTAGCTGTTGTAGATCATAAGTCTACCTATGAAGAAGAT
AGAAAGGACAAGGACATTCCTGTGCAGGAAGAAGTTGAACCTGCCCCAGTTAGGAGGATTCTGAAATTCAGTGCTCCAGA
ATGGCCCTACATGCTGGTAGGGTCTGTGGGTGCAGCTGTGAACGGGACAGTCACACCCTTGTATGCCTTTTTATTCAGCC
AGATTCTTGGGACTTTTTCAATTCCTGATAAAGAGGAACAAAGGTCACAGATCAATGGTGTGTGCCTACTTTTTGTAGCA
ATGGGCTGTGTATCTCTTTTCACCCAATTTCTACAGGATATGCCTTTGCTAAATCTGGGGAGCTCCTAACAAAAAGGCT

ACGTAAATTTGGTTTCAGGGCAATGCTGGGGCAAGATATTGCCTGGTTTGATGACCTCAGAAATAGCCCTGGAGCATTGA
CAACAAGACTTGCTACAGATGCTTCCCAAGTTCAAGGGGCTGCCGGCTCTCAGATCGGGATGATAGTCAATTCCTTCACT
AACGTCACTGTGGCCATGATCATTGCCTTCTCCTTTAGCTGGAAGCTGAGCCTGGTCATCTTGTGCTTCTTCCCCTTCTT
GGCTTTATCAGGAGCCACACAGACCAGGATGTTGACAGGATTTGCCTCTCGAGATAAGCAGGCCCTGGAGATGGTGGGAC
AGATTACAAATGAAGCCCTCAGTAACATCCGCACTGTTGCTGGAATTGGAAAGGAGAGGCGGTTCATTGAAGCACTTGAG
ACTGAGCTGGAGAAGCCCTTCAAGACAGCCATTCAGAAAGCCAATATTTACGGATTCTGCTTTGCCTTTGCCCAGTGCAT
CATGTTTATTGCGAATTCTGCTTCCTACAGATATGGAGGTTACTTAATCTCCAATGAGGGGCTCCATTTCAGCTATGTGT
TCAGGGTGATCTCTGCAGTTGTACTGAGTGCAACAGCTCTTGGAAGAGCCTTCTCTTACACCCCAAGTTATGCAAAAGCT
AAAATATCAGCTGCACGCTTTTTTCAACTGCTGGACCGACAACCCCCAATCAGTGTATACAATACTGCAGGTGAAAAATG
GGACAACTTCCAGGGGAAGATTGATTTTGTTGATTGTAAATTTACATATCCTTCTCGACCTGACTCGCAAGTTCTGAATG
GTCTCTCAGTGTCGATTAGTCCAGGGCAGACACTGGCGTTTGTTGGGAGCAGTGGATGTGGCAAAAGCACTAGCATTCAG
CTGTTGGAACGTTTCTATGATCCTGATCAAGGGAAGGTGATGATAGATGGTCATGACAGCAAAAAAGTAAATGTCCAGTT
CCTCCGCTCAAACATTGGAATTGTTTCCCAGGAACCAGTGTTGTTTGCCTGTAGCATAATGGACAATATCAAGTATGGAG
ACAACACCAAAGAAATTCCCATGGAAAGAGTCATAGCAGCTGCAAAACAGGCTCAGCTGCATGATTTTGTCATGTCACTC
CCAGAGAAATATGAAACTAACGTTGGGTCCCAGGGGTCTCAACTCTCTAGAGGGGAGAAACAACGCATTGCTATTGCTCG
GGCCATTGTACGAGATCCTAAAATCTTGCTACTAGATGAAGCCACTTCTGCCTTAGACACAGAAAGTGAAAAGACGGTGC
AGGTTGCTCTAGACAAAGCCAGAGAGGGTCGGACCTGCATTGTCATTGCCCATCGCTTGTCCACCATCCAGAACGCGGAT
ATCATTGCTGTCATGGCACAGGGGGTGGTGATTGAAAAGGGGACCCATGAAGAACTGATGGCCCAAAAAGGAGCCTACTA
CAAACTAGTCACCACTGGATCCCCCATCAGTTGAGACCCAGCTTTCTTGTACAAAGTGGTGATTGGTACCGAGCTCGGAT
CCACTAGTCCAGTGTGGTGGAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCC
GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAA
GGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
GGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG
CTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGT
GTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCT
CGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACC
TCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACG
TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGA
TTTATAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCT
GTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCTCAAT
TAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGC
AACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGAC
TAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGA
GGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCA
TCGGCATAGTATATCGGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGT
GCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACG
ACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTG
GCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTC
CGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGC
ACTTCGTGGCCGAGGAGCAGGACTGACACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTC
GGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTT
GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATT
CTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTG
TAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACC

TGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGC
TGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG
TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG
TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA
CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA
GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCA
GCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC
TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGT
AGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC
TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG
GTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC
GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC
GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGA
TCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG
GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA
GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAC
GTC

Fig. 11

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTAT
CTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATT
GATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAA
CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
TGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAA
TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
GGAGACCCAAGCTGGCTAGCAGTCCAGGAATCATGCTGGAGAAGTTCTGCAACTCTACTTTTTGGAATTCCTCATTCCTG
GACAGTCCGGAGGCAGACCTGCCACTTTGTTTTGAGCAAACTGTTCTGGTGTGGATTCCCTTGGGCTTCCTATGGCTCCT
GGCCCCCTGGCAGCTTCTCCACGTGTATAAATCCAGGACCAAGAGATCCTCTACCACCAAACTCTATCTTGCTAAGCAGG
TATTCGTTGGTTTTCTTCTTATTCTAGCAGCCATAGAGCTGGCCCTTGTACTCACAGAAGACTCTGGACAAGCCACAGTC
CCTGCTGTTCGATATACCAATCCAAGCCTCTACCTAGGCACATGGCTCCTGGTTTTGCTGATCCAATACAGCAGACAATG

GTGTGTACAGAAAAACTCCTGGTTCCTGTCCCTATTCTGGATTCTCTCGATACTCTGTGGCACTTTCCAATTTCAGACTC
TGATCCGGACACTCTTACAGGGTGACAATTCTAATCTAGCCTACTCCTGCCTGTTCTTCATCTCCTACGGATTCCAGATC
CTGATCCTGATCTTTTCAGCATTTTCAGAAAATAATGAGTCATCAAATAATCCATCATCCATAGCTTCATTCCTGAGTAG
CATTACCTACAGCTGGTATGACAGCATCATTCTGAAAGGCTACAAGCGTCCTCTGACACTCGAGGATGTCTGGGAAGTTG
ATGAAGAGATGAAAACCAAGACATTAGTGAGCAAGTTTGAAACGCACATGAAGAGAGAGCTGCAGAAAGCCAGGCGGGCA
CTCCAGAGACGGCAGGAGAAGAGCTCCCAGCAGAACTCTGGAGCCAGGCTGCCTGGCTTGAACAAGAATCAGAGTCAAAG
CCAAGATGCCCTTGTCCTGGAAGATGTTGAAAAGAAAAAAAAGAAGTCTGGGACCAAAAAAGATGTTCCAAAATCCTGGT
TGATGAAGGCTCTGTTCAAAACTTTCTACATGGTGCTCCTGAAATCATTCCTACTGAAGCTAGTGAATGACATCTTCACG
TTTGTGAGTCCTCAGCTGCTGAAATTGCTGATCTCCTTTGCAAGTGACCGTGACACATATTTGTGGATTGGATATCTCTG
TGCAATCCTCTTATTCACTGCGGCTCTCATTCAGTCTTTCTGCCTTCAGTGTTATTTCCAACTGTGCTTCAAGCTGGGTG
TAAAAGTACGGACAGCTATCATGGCTTCTGTATATAAGAAGGCATTGACCCTATCCAACTTGGCCAGGAAGGAGTACACC
GTTGGAGAAACAGTGAACCTGATGTCTGTGGATGCCCAGAAGCTCATGGATGTGACCAACTTCATGCACATGCTGTGGTC
AAGTGTTCTACAGATTGTCTTATCTATCTTCTTCCTATGGAGAGAGTTGGGACCCTCAGTCTTAGCAGGTGTTGGGGTGA
TGGTGCTTGTAATCCCAATTAATGCGATACTGTCCACCAAGAGTAAGACCATTCAGGTCAAAAATATGAAGAATAAAGAC
AAACGTTTAAAGATCATGAATGAGATTCTTAGTGGAATCAAGATCCTGAAATATTTTGCCTGGGAACCTTCATTCAGAGA
CCAAGTACAAAACCTCCGGAAGAAAGAGCTCAAGAACCTGCTGGCCTTTAGTCAACTACAGTGTGTAGTAATATTCGTCT
TCCAGTTAACTCCAGTCCTGGTATCTGTGGTCACATTTTCTGTTTATGTCCTGGTGGATAGCAACAATATTTTGGATGCA
CAAAAGGCCTTCACCTCCATTACCCTCTTCAATATCCTGCGCTTTCCCCTGAGCATGCTTCCCATGATGATCTCCTCCAT
GCTCCAGGCCAGTGTTTCCACAGAGCGGCTAGAGAAGTACTTGGGAGGGGATGACTTGGACACATCTGCCATTCGACATG
ACTGCAATTTTGACAAAGCCATGCAGTTTTCTGAGGCCTCCTTTACCTGGGAACATGATTCGGAAGCCACAGTCCGAGAT
GTGAACCTGGACATTATGGCAGGCCAACTTGTGGCTGTGATAGGCCCTGTCGGCTCTGGGAAATCCTCCTTGATATCAGC
CATGCTGGGAGAAATGGAAAATGTCCACGGGCACATCACCATCAAGGGCACCACTGCCTATGTCCCACAGCAGTCCTGGA
TTCAGAATGGCACCATAAAGGACAACATCCTTTTTGGAACAGAGTTTAATGAAAAGAGGTACCAGCAAGTACTGGAGGCC
TGTGCTCTCCTCCCAGACTTGGAAATGCTGCCTGGAGGAGATTTGGCTGAGATTGGAGAGAAGGGTATAAATCTTAGTGG
GGGTCAGAAGCAGCGGATCAGCCTGGCCAGAGCTACCTACCAAAATTTAGACATCTATCTTCTAGATGACCCCCTGTCTG
CAGTGGATGCTCATGTAGGAAAACATATTTTTAATAAGGTCTTGGGCCCCAATGGCCTGTTGAAAGGCAAGACTCGACTC
TTGGTTACACATAGCATGCACTTTCTTCCTCAAGTGGATGAGATTGTAGTTCTGGGGAATGGAACAATTGTAGAGAAAGG
ATCCTACAGTGCTCTCCTGGCCAAAAAAGGAGAGTTTGCTAAGAATCTGAAGACATTTCTAAGACATACAGGCCCTGAAG
AGGAAGCCACAGTCCATGATGGCAGTGAAGAAGAAGACGATGACTATGGGCTGATATCCAGTGTGGAAGAGATCCCCGAA
GATGCAGCCTCCATAACCATGAGAAGAGAGAACAGCTTTCGTCGAACACTTAGCCGCAGTTCTAGGTCCAATGGCAGGCA
TCTGAAGTCCCTGAGAAACTCCTTGAAAACTCGGAATGTGAATAGCCTGAAGGAAGACGAAGAACTAGTGAAAGGACAAA
AACTAATTAAGAAGGAATTCATAGAAACTGGAAAGGTGAAGTTCTCCATCTACCTGGAGTACCTACAAGCAATAGGATTG
TTTTCGATATTCTTCATCATCCTTGCGTTTGTGATGAATTCTGTGGCTTTTATTGGATCCAACCTCTGGCTCAGTGCTTG
GACCAGTGACTCTAAAATCTTCAATAGCACCGACTATCCAGCATCTCAGAGGGACATGAGAGTTGGAGTCTACGGAGCTC
TGGGATTAGCCCAAGGTATATTTGTGTTCATAGCACATTTCTGGAGTGCCTTTGGTTTCGTCCATGCATCAAATATCTTG
CACAAGCAACTGCTGAACAATATCCTTCGAGCACCTATGAGATTTTTTGACACAACACCCACAGGCCGGATTGTGAACAG
GTTTGCCGGCGATATTTCCACAGTGGATGACACCCTGCCTCAGTCCTTGCGCAGCTGGATTACATGCTTCCTGGGGATAA
TCAGCACCCTTGTCATGATCTGCATGGCCACTCCTGTCTTCACCATCATCGTCATTCCTCTTGGCATTATTTATGTATCT
GTTCAGATGTTTTATGTGTCTACCTCCCGCCAGCTGAGGCGTCTGGACTCTGTCACCAGGTCCCCAATCTACTCTCACTT
CAGCGAGACCGTATCAGGTTTGCCAGTTATCCGTGCCTTTGAGCACCAGCAGCGATTTCTGAAACACAATGAGGTGAGGA
TTGACACCAACCAGAAATGTGTCTTTTCCTGGATCACCTCCAACAGGTGGCTTGCAATTCGCCTGGAGCTGGTTGGGAAC
CTGACTGTCTTCTTTTCAGCCTTGATGATGGTTATTTATAGAGATACCCTAAGTGGGGACACTGTTGGCTTTGTTCTGTC
CAATGCACTCAATATCACACAAACCCTGAACTGGCTGGTGAGGATGACATCAGAAATAGAGACCAACATTGTGGCTGTTG
AGCGAATAACTGAGTACACAAAAGTGGAAAATGAGGCACCCTGGGTGACTGATAAGAGGCCTCCGCCAGATTGGCCCAGC
AAAGGCAAGATCCAGTTTAACAACTACCAAGTGCGGTACCGACCTGAGCTGGATCTGGTCCTCAGAGGGATCACTTGTGA

CATTGGTAGCATGGAGAAGATTGGTGTGGTGGGCAGGACAGGAGCTGGAAAGTCATCCCTCACAAACTGCCTCTTCAGAA
TCTTAGAGGCTGCCGGTGGTCAGATTATCATTGATGGAGTAGATATTGCTTCCATTGGGCTCCACGACCTCCGAGAGAAG
CTGACCATCATCCCCCAGGACCCCATCCTGTTCTCTGGAAGCCTGAGGATGAATCTCGACCCTTTCAACAACTACTCAGA
TGAGGAGATTTGGAAGGCCTTGGAGCTGGCTCACCTCAAGTCTTTTGTGGCCAGCCTGCAACTTGGGTTATCCCACGAAG
TGACAGAGGCTGGTGGCAACCTGAGCATAGGCCAGAGGCAGCTGCTGTGCCTGGGCAGGGCTCTGCTTCGGAAATCCAAG
ATCCTGGTCCTGGATGAGGCCACTGCTGCGGTGGATCTAGAGACAGACAACCTCATTCAGACGACCATCCAAAACGAGTT
CGCCCACTGCACAGTGATCACCATCGCCCACAGGCTGCACACCATCATGGACAGTGACAAGGTAATGGTCCTAGACAACG
GGAAGATTATAGAGTGCGGCAGCCCTGAAGAACTGCTACAAATCCCTGGACCCTTTTACTTTATGGCTAAGGAAGCTGGC
ATTGAGAATGTGAACAGCACAAAATTCTAGCTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTC
TGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTC
TAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCT
AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTG
GGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGT
CAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGG
TGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTG
GACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGGGGATTTCG
GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGT
GGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGT
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCG
CCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGC
CGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCC
CGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTAT
AATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGCGACGTCGCCGGA
GCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGA
CGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGG
ACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGC
GAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTG
ACACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT
GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTAC
AAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGT
GAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAAT
CGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCG
TTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCC
CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGA
AGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT
CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAA

AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

SEQUENCE LISTING

<110> Aventis Pharma Deutschland GmbH

<120> Dual-transfected cell llines as in vitro screening tools for pharmaceutical compound profiling: A model for hepatobiliary elimination

<130> DEAV 2004/0025

<160> 6

<170> PatentIn version 3.2

<210> 1
<211> 1061
<212> DNA
<213> Homo sapiens

<400> 1
```
caggaggatg gaggcccaca acgcgtctgc cccattcaac ttcaccctgc cacccaactt      60
[illegible]ggcaagcgc cccacagacc tggcactgag cgtcatcctg gtgttcatgt tgttcttcat     120
[illegible]tgctctcg ctgggctgca ccatggagtt cagcaagatc aaggctcact tatggaagcc     180
taaagggctg gccatcgccc tggtggcaca gtatggcatc atgcccctca cggcctttgt     240
gctgggcaag gtcttccggc tgaagaacat tgaggcactg gccatcttgg tctgtggctg     300
ctcacctgga gggaacctgt ccaatgtctt cagtctggcc atgaaggggg acatgaacct     360
cagcattgtg atgaccacct gctccacctt ctgtgccctt ggcatgatgc ctctcctcct     420
gtacatctac tccaggggga tctatgatgg ggacctgaag gacaaggtgc cctataaagg     480
catcgtgata tcactggtcc tggttctcat tccttgcacc atagggatcg tcctcaaatc     540
taaacggcca caatacatgc gctatgtcat caagggaggg atgatcatca ttctcttgtg     600
cagtgtggcc gtcacagttc tctctgccat caatgtgggg aagagcatca tgtttgccat     660
gacaccactc ttgattgcca cctcctccct gatgccttct attggctttc tgctgggtta     720
tgttctctct gctctcttct gcctcaatgg acggtgcaga cgcactgtca gcatggagac     780
[illegible]gatgccaa aatgtccaac tctgttccac catcctcaat gtggcctttc cacctgaagt     840
cattggacca cttttcttct ttcccctcct ctacatgatt ttccagcttg gagaagggct     900
tctcctcatt gccatatttt ggtgctatga gaaattcaag actcccaagg ataaaacaaa     960
aatgatctac acagctgcca caactgaaga aacaattcca ggagctctgg gaaatggcac    1020
ctacaaaggg gaggactgct cccettgcac agcctagccc t                       1061
```

<210> 2
<211> 3966
<212> DNA
<213> Homo sapiens

<400> 2
```
atgtctgact cagtaattct tcgaagtata aagaaatttg gagaggagaa tgatggtttt      60
gagtcagata aatcatataa taatgataag aaatcaaggt tacaagatga gaagaaaggt     120
gatggcgtta gagttggctt ctttcaattg tttcggtttt cttcatcaac tgacatttgg     180
```

ctgatgtttg tgggaagttt gtgtgcattt ctccatggaa tagcccagcc aggcgtgcta 240
ctcattttg gcacaatgac agatgttttt attgactacg acgttgagtt acaagaactc 300
cagattccag gaaaagcatg tgtgaataac accattgtat ggactaacag ttccctcaac 360
cagaacatga caaatggaac acgttgtggg ttgctgaaca tcgagagcga aatgatcaaa 420
tttgccagtt actatgctgg aattgctgtc gcagtactta tcacaggata tattcaaata 480
tgcttttggg tcattgccgc agctcgtcag atacagaaaa tgagaaaatt ttactttagg 540
agaataatga gaatggaaat agggtggttt gactgcaatt cagtggggga gctgaataca 600
agattctctg atgatattaa taaaatcaat gatgccatag ctgaccaaat ggcccttttc 660
attcagcgca tgacctcgac catctgtggt ttcctgttgg gatttttcag ggttggaaa 720
ctgaccttgg ttattatttc tgtcagccct ctcattggga ttggagcagc caccattggt 780
ctgagtgtgt ccaagtttac ggactatgag ctgaaggcct atgccaaagc aggggtggtg 840
ttgatgaag tcatttcatc aatgagaaca gtggctgctt ttggtggtga gaaaagagag 900
gttgaaaggt atgagaaaaa tcttgtgttc gcccagcgtt ggggaattag aaaaggaata 960
gtgatgggat tctttactgg attcgtgtgg tgtctcatct tttgtgtta tgcactggcc 1020
ttctggtacg gctccacact tgtcctggat gaaggagaat atacaccagg aaccccttgtc 1080
cagattttcc tcagtgtcat agtaggagct ttaaatcttg gcaatgcctc tccttgtttg 1140
gaagcctttg caactggacg tgcagcagcc accagcattt ttgagacaat agacaggaaa 1200
cccatcattg actgcatgtc agaagatggt tacaagttgg atcgaatcaa gggtgaaatt 1260
gaattccata atgtgacctt ccattatcct tccagaccag aggtgaagat tctaaatgac 1320
ctcaacatgg ccattaaacc aggggaaatg acagctctgg taggacccag tggagctgga 1380
aaaagtacag cactgcaact cattcagcga ttctatgacc cctgtgaagg aatggtgacc 1440
gtggatggcc atgacattcg ctctcttaac attcagtggc ttagagatca gattgggata 1500
tggagcaag agccagttct gttctctacc accattgcag aaaatattcg ctatggcaga 1560
aagatgcaa caatggaaga catagtccaa gctgccaagg aggccaatgc ctacaacttc 1620
atcatggacc tgccacagca atttgacacc cttgttggag aaggaggagg ccagatgagt 1680
ggtggccaga aacaaagggt agctatcgcc agagccctca tccgaaatcc caagattctg 1740
cttttggaca tggccacctc agctctggac aatgagagtg aagccatggt gcaagaagtg 1800
ctgagtaaga ttcagcatgg gcacacaatc atttcagttg ctcatcgctt gtctacggtc 1860
agagctgcag ataccatcat tggttttgaa catggcactg cagtggaaag agggacccat 1920
gaagaattac tggaaaggaa aggtgtttac ttcactctag tgactttgca aagccaggga 1980
aatcaagctc ttaatgaaga ggacataaag gatgcaactg aagatgacat gcttgcgagg 2040
acctttagca gagggagcta ccaggatagt ttaagggctt ccatccggca acgctccaag 2100
tctcagcttt cttacctggt gcacgaacct ccattagctg ttgtagatca taagtctacc 2160
tatgaagaag atagaaagga caaggacatt cctgtgcagg aagaagttga acctgcccca 2220

gttaggagga ttctgaaatt cagtgctcca gaatggccct acatgctggt agggtctgtg 2280
ggtgcagctg tgaacgggac agtcacaccc ttgtatgcct ttttattcag ccagattctt 2340
gggacttttt caattcctga taaagaggaa caaaggtcac agatcaatgg tgtgtgccta 2400
ctttttgtag caatgggctg tgtatctctt ttcacccaat ttctacaggg atatgccttt 2460
gctaaatctg gggagctcct aacaaaaagg ctacgtaaat ttggtttcag ggcaatgctg 2520
gggcaagata ttgcctggtt tgatgacctc agaaatagcc ctggagcatt gacaacaaga 2580
cttgctacag atgcttccca agttcaaggg gctgccggct ctcagatcgg gatgatagtc 2640
aattccttca ctaacgtcac tgtggccatg atcattgcct tctcctttag ctggaagctg 2700
agcctggtca tcttgtgctt cttccccttc ttggctttat caggagccac acagaccagg 2760
atgttgacag gatttgcctc tcgagataag caggccctgg agatggtggg acagattaca 2820
aatgaagccc tcagtaacat ccgcactgtt gctggaattg gaaaggagag gcggttcatt 2880
[illegible]agcacttg agactgagct ggagaagccc ttcaagacag ccattcagaa agccaatatt 2940
[illegible]acggattct gctttgcctt tgcccagtgc atcatgttta ttgcgaattc tgcttcctac 3000
agatatggag gttacttaat ctccaatgag ggctccattt tcagctatgt gttcagggtg 3060
atctctgcag ttgtactgag tgcaacagct cttggaagag ccttctctta caccccaagt 3120
tatgcaaaag ctaaaatatc agctgcacgc ttttttcaac tgctggaccg acaacccccca 3180
atcagtgtat acaatactgc aggtgaaaaa tgggacaact tccaggggaa gattgatttt 3240
gttgattgta aatttacata tccttctcga cctgactcgc aagttctgaa tggtctctca 3300
gtgtcgatta gtccagggca gacactggcg tttgttggga gcagtggatg tggcaaaagc 3360
actagcattc agctgttgga acgtttctat gatcctgatc aagggaaggt gatgatagat 3420
ggtcatgaca gcaaaaaagt aaatgtccag ttcctccgct caaacattgg aattgtttcc 3480
caggaaccag tgttgtttgc ctgtagcata atggacaata tcaagtatgg agacaacacc 3540
[illegible]aagaaattc ccatggaaag agtcatagca gctgcaaaac aggctcagct gcatgatttt 3600
[illegible]catgtcac tcccagagaa atatgaaact aacgttgggt cccaggggtc tcaactctct 3660
agaggggaga aacaacgcat tgctattgct cgggccattg tacgagatcc taaaatcttg 3720
ctactagatg aagccacttc tgccttagac acagaaagtg aaaagacggt gcaggttgct 3780
ctagacaaag ccagagaggg tcggacctgc attgtcattg cccatcgctt gtccaccatc 3840
cagaacgcgg atatcattgc tgtcatggca caggggtgg tgattgaaaa ggggacccat 3900
gaagaactga tggcccaaaa aggagcctac tacaaactag tcaccactgg atcccccatc 3960
agttga 3966

<210> 3
<211> 4650
<212> DNA
<213> Homo sapiens

<400> 3

```
agtccaggaa tcatgctgga gaagttctgc aactctactt tttggaattc ctcattcctg     60
gacagtccgg aggcagacct gccactttgt tttgagcaaa ctgttctggt gtggattccc    120
ttgggcttcc tatggctcct ggcccctgg cagcttctcc acgtgtataa atccaggacc     180
aagagatcct ctaccaccaa actctatctt gctaagcagg tattcgttgg ttttcttctt    240
attctagcag ccatagagct ggcccttgta ctcacagaag actctggaca agccacagtc    300
cctgctgttc gatataccaa tccaagcctc tacctaggca catggctcct ggttttgctg    360
atccaataca gcagacaatg gtgtgtacag aaaaactcct ggttcctgtc cctattctgg    420
attctctcga tactctgtgg cactttccaa tttcagactc tgatccggac actcttacag    480
ggtgacaatt ctaatctagc ctactcctgc ctgttcttca tctcctacgg attccagatc    540
ctgatcctga tcttttcagc attttcagaa aataatgagt catcaaataa tccatcatcc    600
atagcttcat tcctgagtag cattacctac agctggtatg acagcatcat tctgaaaggc    660
tacaagcgtc ctctgacact cgaggatgtc tgggaagttg atgaagagat gaaaaccaag    720
[illegible]tattagtga gcaagtttga aacgcacatg aagagagagc tgcagaaagc caggcgggca    780
ctccagagac ggcaggagaa gagctcccag cagaactctg gagccaggct gcctggcttg    840
aacaagaatc agagtcaaag ccaagatgcc cttgtcctgg aagatgttga aaagaaaaaa    900
aagaagtctg ggaccaaaaa agatgttcca aaatcctggt tgatgaaggc tctgttcaaa    960
actttctaca tggtgctcct gaaatcattc ctactgaagc tagtgaatga catcttcacg   1020
tttgtgagtc ctcagctgct gaaattgctg atctcctttg caagtgaccg tgacacatat   1080
ttgtggattg gatatctctg tgcaatcctc ttattcactg cggctctcat tcagtctttc   1140
tgccttcagt gttatttcca actgtgcttc aagctgggtg taaaagtacg gacagctatc   1200
atggcttctg tatataagaa ggcattgacc ctatccaact tggccaggaa ggagtacacc   1260
gttggagaaa cagtgaacct gatgtctgtg gatgcccaga agctcatgga tgtgaccaac   1320
ttcatgcaca tgctgtggtc aagtgttcta cagattgtct tatctatctt cttcctatgg   1380
[illegible]agagttgg gaccctcagt cttagcaggt gttggggtga tggtgcttgt aatcccaatt   1440
aatgcgatac tgtccaccaa gagtaagacc attcaggtca aaaatatgaa gaataaagac   1500
aaacgtttaa agatcatgaa tgagattctt agtggaatca agatcctgaa atattttgcc   1560
tgggaacctt cattcagaga ccaagtacaa aacctccgga agaaagagct caagaacctg   1620
ctggccttta gtcaactaca gtgtgtagta atattcgtct tccagttaac tccagtcctg   1680
gtatctgtgg tcacattttc tgtttatgtc ctggtggata gcaacaatat tttggatgca   1740
caaaaggcct tcacctccat taccctcttc aatatcctgc gctttcccct gagcatgctt   1800
cccatgatga tctcctccat gctccaggcc agtgtttcca cagagcggct agagaagtac   1860
ttgggagggg atgacttgga cacatctgcc attcgacatg actgcaattt tgacaaagcc   1920
atgcagtttt ctgaggcctc ctttacctgg gaacatgatt cggaagccac agtccgagat   1980
gtgaacctgg acattatggc aggccaactt gtggctgtga taggccctgt cggctctggg   2040
```

```
aaatcctcct tgatatcagc catgctggga gaaatggaaa atgtccacgg gcacatcacc   2100
atcaagggca ccactgccta tgtcccacag cagtcctgga ttcagaatgg caccataaag   2160
gacaacatcc tttttggaac agagtttaat gaaaagaggt accagcaagt actggaggcc   2220
tgtgctctcc tcccagactt ggaaatgctg cctggaggag atttggctga gattggagag   2280
aagggtataa atcttagtgg gggtcagaag cagcggatca gcctggccag agctacctac   2340
caaaatttag acatctatct tctagatgac cccctgtctg cagtggatgc tcatgtagga   2400
aaacatattt ttaataaggt cttgggcccc aatggcctgt tgaaaggcaa gactcgactc   2460
ttggttacac atagcatgca ctttcttcct caagtggatg agattgtagt tctggggaat   2520
ggaacaattg tagagaaagg atcctacagt gctctcctgg ccaaaaaagg agagtttgct   2580
aagaatctga agacatttct aagacataca ggccctgaag aggaagccac agtccatgat   2640
ggcagtgaag aagaagacga tgactatggg ctgatatcca gtgtggaaga gatccccgaa   2700
gatgcagcct ccataaccat gagaagagag aacagctttc gtcgaacact tagccgcagt   2760
[illegible]taggtcca atggcaggca tctgaagtcc ctgagaaact ccttgaaaac tcggaatgtg   2820
aatagcctga aggaagacga agaactagtg aaaggacaaa aactaattaa gaaggaattc   2880
atagaaactg gaaaggtgaa gttctccatc tacctggagt acctacaagc aataggattg   2940
ttttcgatat tcttcatcat ccttgcgttt gtgatgaatt ctgtggcttt tattggatcc   3000
aacctctggc tcagtgcttg gaccagtgac tctaaaatct tcaatagcac cgactatcca   3060
gcatctcaga gggacatgag agttggagtc tacggagctc tgggattagc ccaaggtata   3120
tttgtgttca tagcacattt ctggagtgcc tttggtttcg tccatgcatc aaatatcttg   3180
cacaagcaac tgctgaacaa tatccttcga gcacctatga gatttttttga cacaacaccc   3240
acaggccgga ttgtgaacag gtttgccggc gatatttcca cagtggatga caccctgcct   3300
cagtccttgc gcagctggat tacatgcttc ctggggataa tcagcaccct tgtcatgatc   3360
tgcatggcca ctcctgtctt caccatcatc gtcattcctc ttggcattat ttatgtatct   3420
[illegible]tcagatgt tttatgtgtc tacctcccgc cagctgaggc gtctggactc tgtcaccagg   3480
tccccaatct actctcactt cagcgagacc gtatcaggtt tgccagttat ccgtgccttt   3540
gagcaccagc agcgatttct gaaacacaat gaggtgagga ttgacaccaa ccagaaatgt   3600
gtcttttcct ggatcacctc caacaggtgg cttgcaattc gcctggagct ggttgggaac   3660
ctgactgtct tcttttcagc cttgatgatg gttatttata gagataccct aagtggggac   3720
actgttggct ttgttctgtc caatgcactc aatatcacac aaaccctgaa ctggctggtg   3780
aggatgacat cagaaataga gaccaacatt gtggctgttg agcgaataac tgagtacaca   3840
aaagtggaaa atgaggcacc ctgggtgact gataagaggc ctccgccaga ttggcccagc   3900
aaaggcaaga tccagtttaa caactaccaa gtgcggtacc gacctgagct ggatctggtc   3960
ctcagaggga tcacttgtga cattggtagc atggagaaga ttggtgtggt gggcaggaca   4020
ggagctggaa agtcatccct cacaaactgc ctcttcagaa tcttagaggc tgccggtggt   4080
```

cagattatca ttgatggagt agatattgct tccattgggc tccacgacct ccgagagaag 4140

ctgaccatca tcccccagga ccccatcctg ttctctggaa gcctgaggat gaatctcgac 4200

cctttcaaca actactcaga tgaggagatt tggaaggcct tggagctggc tcacctcaag 4260

tcttttgtgg ccagcctgca acttgggtta tcccacgaag tgacagaggc tggtggcaac 4320

ctgagcatag gccagaggca gctgctgtgc ctgggcaggg ctctgcttcg gaaatccaag 4380

atcctggtcc tggatgaggc cactgctgcg gtggatctag agacagacaa cctcattcag 4440

acgaccatcc aaaacgagtt cgcccactgc acagtgatca ccatcgccca caggctgcac 4500

accatcatgg acagtgacaa ggtaatggtc ctagacaacg ggaagattat agagtgcggc 4560

agccctgaag aactgctaca aatccctgga cccttttact ttatggctaa ggaagctggc 4620

attgagaatg tgaacagcac aaaattctag 4650

<210> 4
<211> 6433
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 4
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420

attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480

[illegible]catatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540

atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660

actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720

aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840

ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900

gtttaaactt aagcttggta ccgagctcgg atcccaggag gatggaggcc cacaacgcgt 960

ctgccccatt caacttcacc ctgccaccca actttggcaa gcgccccaca gacctggcac 1020

tgagcgtcat cctggtgttc atgttgttct tcatcatgct ctcgctgggc tgcaccatgg 1080

agttcagcaa gatcaaggct cacttatgga agcctaaagg gctggccatc gccctggtgg 1140

cacagtatgg catcatgccc ctcacggcct ttgtgctggg caaggtcttc cggctgaaga 1200
acattgaggc actggccatc ttggtctgtg gctgctcacc tggagggaac ctgtccaatg 1260
tcttcagtct ggccatgaag gggacatga acctcagcat tgtgatgacc acctgctcca 1320
ccttctgtgc ccttggcatg atgcctctcc tcctgtacat ctactccagg gggatctatg 1380
atggggacct gaaggacaag gtgccctata aaggcatcgt gatatcactg gtcctggttc 1440
tcattccttg caccataggg atcgtcctca aatctaaacg gccacaatac atgcgctatg 1500
tcatcaaggg agggatgatc atcattctct tgtgcagtgt ggccgtcaca gttctctctg 1560
ccatcaatgt ggggaagagc atcatgtttg ccatgacacc actcttgatt gccacctcct 1620
ccctgatgcc ttctattggc tttctgctgg gttatgttct ctctgctctc ttctgcctca 1680
atggacggtg cagacgcact gtcagcatgg agactggatg ccaaaatgtc caactctgtt 1740
ccaccatcct caatgtggcc tttccacctg aagtcattgg accacttttc ttctttcccc 1800
tcctctacat gattttccag cttggagaag ggcttctcct cattgccata ttttggtgct 1860
[illegible]gagaaatt caagactccc aaggataaaa caaaaatgat ctacacagct gccacaactg 1920
aagaaacaat tccaggagct ctgggaaatg gcacctacaa aggggaggac tgctcccctt 1980
gcacagccta gcccttctag agggcccgtt taaacccgct gatcagcctc gactgtgcct 2040
tctagttgcc agccatctgt tgtttgcccc tcccccgtgc cttccttgac cctggaaggt 2100
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg 2160
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 2220
aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga aagaaccagc 2280
tggggctcta gggggtatcc ccacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg 2340
gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcccgc tcctttcgct 2400
ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg 2460
ctccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaaa acttgattag 2520
[illegible]tgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg 2580
gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc 2640
tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat 2700
gagctgattt aacaaaaatt taacgcgaat taattctgtg gaatgtgtgt cagttagggt 2760
gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat ctcaattagt 2820
cagcaaccag gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc 2880
atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccatcccg cccctaactc 2940
cgcccagttc cgcccattct ccgccccatg gctgactaat tttttttatt tatgcagagg 3000
ccgaggccgc ctctgcctct gagctattcc agaagtagtg aggaggcttt tttggaggcc 3060
taggcttttg caaaaagctc ccgggagctt gtatatccat tttcggatct gatcaagaga 3120
caggatgagg atcgtttcgc atgattgaac aagatggatt gcacgcaggt tctccggccg 3180

cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc tgctctgatg 3240
ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct ttttgtcaag accgacctgt 3300
ccggtgccct gaatgaactg caggacgagg cagcgcggct atcgtggctg gccacgacgg 3360
gcgttccttg cgcagctgtg ctcgacgttg tcactgaagc gggaagggac tggctgctat 3420
tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat 3480
ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc tgcccattcg 3540
accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg 3600
atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc 3660
tcaaggcgcg catgcccgac ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc 3720
cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg 3780
tggcggaccg ctatcaggac atagcgttgg ctacccgtga tattgctgaa gagcttggcg 3840
gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca 3900
tcgccttcta tcgccttctt gacgagttct tctgagcggg actctggggt tcgaaatgac 3960
cgaccaagcg acgcccaacc tgccatcacg agatttcgat tccaccgccg ccttctatga 4020
aaggttgggc ttcggaatcg ttttccggga cgccggctgg atgatcctcc agcgcgggga 4080
tctcatgctg gagttcttcg cccaccccaa cttgtttatt gcagcttata atggttacaa 4140
ataaagcaat agcatcacaa atttcacaaa taaagcattt ttttcactgc attctagttg 4200
tggtttgtcc aaactcatca atgtatctta tcatgtctgt ataccgtcga cctctagcta 4260
gagcttggcg taatcatggt catagctgtt tcctgtgtga aattgttatc cgctcacaat 4320
tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag 4380
ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg 4440
ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc 4500
ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc 4560
agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa 4620
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 4680
tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 4740
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 4800
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 4860
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 4920
caagctgggc tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa 4980
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 5040
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 5100
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac 5160
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggttt 5220

ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 5280

cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 5340

gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 5400

aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 5460

acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 5520

gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 5580

cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 5640

cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 5700

tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat 5760

cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 5820

gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 5880

cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 5940

ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa 6000

gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga 6060

taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg 6120

gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 6180

acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 6240

aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 6300

cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6360

atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6420

gccacctgac gtc 6433

<210> 5
<211> 9043
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 5
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420

attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480

```
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc     900
gtttaaactt aagctatcac aagtttgtac aaaaaagcag gcttaggaat gtctgactca     960
gtaattcttc gaagtataaa gaaatttgga gaggagaatg atggttttga gtcagataaa    1020
tcatataata atgataagaa atcaaggtta caagatgaga agaaaggtga tggcgttaga    1080
gttggcttct ttcaattgtt tcggttttct tcatcaactg acatttggct gatgtttgtg    1140
ggaagtttgt gtgcatttct ccatggaata gcccagccag gcgtgctact catttttggc    1200
[illegible]taatgacag atgtttttat tgactacgac gttgagttac aagaactcca gattccagga    1260
aaagcatgtg tgaataacac cattgtatgg actaacagtt ccctcaacca gaacatgaca    1320
aatggaacac gttgtgggtt gctgaacatc gagagcgaaa tgatcaaatt tgccagttac    1380
tatgctggaa ttgctgtcgc agtacttatc acaggatata ttcaaatatg cttttgggtc    1440
attgccgcag ctcgtcagat acagaaaatg agaaaatttt actttaggag aataatgaga    1500
atggaaatag ggtggtttga ctgcaattca gtgggggagc tgaatacaag attctctgat    1560
gatattaata aaatcaatga tgccatagct gaccaaatgg cccttttcat tcagcgcatg    1620
acctcgacca tctgtggttt cctgttggga tttttcaggg gttggaaact gaccttggtt    1680
attatttctg tcagccctct cattgggatt ggagcagcca ccattggtct gagtgtgtcc    1740
aagtttacgg actatgagct gaaggcctat gccaaagcag gggtggtggc tgatgaagtc    1800
atttcatcaa tgagaacagt ggctgctttt ggtggtgaga aaagagaggt tgaaaggtat    1860
[illegible]gaaaaatc ttgtgttcgc ccagcgttgg ggaattagaa aaggaatagt gatgggattc    1920
cttactggat tcgtgtggtg tctcatcttt ttgtgttatg cactggcctt ctggtacggc    1980
tccacacttg tcctggatga aggagaatat acaccaggaa cccttgtcca gattttcctc    2040
agtgtcatag taggagcttt aaatcttggc aatgcctctc cttgtttgga agcctttgca    2100
actggacgtg cagcagccac cagcattttt gagacaatag acaggaaacc catcattgac    2160
tgcatgtcag aagatggtta caagttggat cgaatcaagg gtgaaattga attccataat    2220
gtgaccttcc attatccttc cagaccagag gtgaagattc taaatgacct caacatggcc    2280
attaaaccag gggaaatgac agctctggta ggacccagtg gagctggaaa aagtacagca    2340
ctgcaactca ttcagcgatt ctatgacccc tgtgaaggaa tggtgaccgt ggatggccat    2400
gacattcgct ctcttaacat tcagtggctt agagatcaga ttgggatagt ggagcaagag    2460
ccagttctgt tctctaccac cattgcagaa aatattcgct atggcagaga agatgcaaca    2520
```

atggaagaca tagtccaagc tgccaaggag gccaatgcct acaacttcat catggacctg 2580
ccacagcaat ttgacaccct tgttggagaa ggaggaggcc agatgagtgg tggccagaaa 2640
caaagggtag ctatcgccag agccctcatc cgaaatccca agattctgct tttggacatg 2700
gccacctcag ctctggacaa tgagagtgaa gccatggtgc aagaagtgct gagtaagatt 2760
cagcatgggc acacaatcat ttcagttgct catcgcttgt ctacggtcag agctgcagat 2820
accatcattg gttttgaaca tggcactgca gtggaaagag ggacccatga agaattactg 2880
gaaaggaaag gtgtttactt cactctagtg actttgcaaa gccagggaaa tcaagctctt 2940
aatgaagagg acataaagga tgcaactgaa gatgacatgc ttgcgaggac ctttagcaga 3000
gggagctacc aggatagttt aagggcttcc atccggcaac gctccaagtc tcagctttct 3060
tacctggtgc acgaacctcc attagctgtt gtagatcata agtctaccta tgaagaagat 3120
agaaaggaca aggacattcc tgtgcaggaa gaagttgaac ctgccccagt taggaggatt 3180
ctgaaattca gtgctccaga atggccctac atgctggtag ggtctgtggg tgcagctgtg 3240
[illegible]cgggacag tcacaccctt gtatgccttt ttattcagcc agattcttgg gactttttca 3300
attcctgata aagaggaaca aaggtcacag atcaatggtg tgtgcctact tttgtagca 3360
atgggctgtg tatctctttt cacccaattt ctacagggat atgcctttgc taaatctggg 3420
gagctcctaa caaaaaggct acgtaaattt ggtttcaggg caatgctggg gcaagatatt 3480
gcctggtttg atgacctcag aaatagccct ggagcattga caacaagact tgctacagat 3540
gcttcccaag ttcaaggggc tgccggctct cagatcggga tgatagtcaa ttccttcact 3600
aacgtcactg tggccatgat cattgccttc tcctttagct ggaagctgag cctggtcatc 3660
ttgtgcttct tccccttctt ggctttatca ggagccacac agaccaggat gttgacagga 3720
tttgcctctc gagataagca ggccctggag atggtgggac agattacaaa tgaagccctc 3780
agtaacatcc gcactgttgc tggaattgga aaggagaggc ggttcattga agcacttgag 3840
actgagctgg agaagccctt caagacagcc attcagaaag ccaatattta cggattctgc 3900
[illegible]tgcctttg cccagtgcat catgtttatt gcgaattctg cttcctacag atatggaggt 3960
cacttaatct ccaatgaggg gctccatttc agctatgtgt tcagggtgat ctctgcagtt 4020
gtactgagtg caacagctct tggaagagcc ttctcttaca ccccaagtta tgcaaaagct 4080
aaaatatcag ctgcacgctt ttttcaactg ctggaccgac aacccccaat cagtgtatac 4140
aatactgcag gtgaaaaatg ggacaacttc caggggaaga ttgattttgt tgattgtaaa 4200
tttacatatc cttctcgacc tgactcgcaa gttctgaatg gtctctcagt gtcgattagt 4260
ccagggcaga cactggcgtt tgttgggagc agtggatgtg gcaaaagcac tagcattcag 4320
ctgttggaac gtttctatga tcctgatcaa gggaaggtga tgatagatgg tcatgacagc 4380
aaaaaagtaa atgtccagtt cctccgctca aacattggaa ttgtttccca ggaaccagtg 4440
ttgtttgcct gtagcataat ggacaatatc aagtatggag acaacaccaa agaaattccc 4500
atggaaagag tcatagcagc tgcaaaacag gctcagctgc atgattttgt catgtcactc 4560

ccagagaaat atgaaactaa cgttgggtcc caggggtctc aactctctag aggggagaaa 4620
caacgcattg ctattgctcg ggccattgta cgagatccta aaatcttgct actagatgaa 4680
gccacttctg ccttagacac agaaagtgaa aagacggtgc aggttgctct agacaaagcc 4740
agagagggtc ggacctgcat tgtcattgcc catcgcttgt ccaccatcca gaacgcggat 4800
atcattgctg tcatggcaca gggggtggtg attgaaaagg ggacccatga agaactgatg 4860
gcccaaaaag gagcctacta caaactagtc accactggat cccccatcag ttgagaccca 4920
gctttcttgt acaaagtggt gattggtacc gagctcggat ccactagtcc agtgtggtgg 4980
aattctgcag atatccagca cagtggcggc cgctcgagtc tagagggccc gtttaaaccc 5040
gctgatcagc ctcgactgtg ccttctagtt gccagccatc tgttgtttgc ccctcccccg 5100
tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa 5160
ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtggggtg gggcaggaca 5220
gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg 5280
cttctgaggc ggaaagaacc agctggggct ctagggggta tccccacgcg ccctgtagcg 5340
gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg 5400
ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc 5460
cccgtcaagc tctaaatcgg ggcatccctt tagggttccg atttagtgct ttacggcacc 5520
tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga 5580
cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa 5640
ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg attttgggga 5700
tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattaattct 5760
gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc tccccaggca ggcagaagta 5820
tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg aaagtcccca ggctccccag 5880
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccatagtc ccgcccctaa 5940
ctccgcccat cccgccccta actccgccca gttccgccca ttctccgccc catggctgac 6000
taattttttt tatttatgca gaggccgagg ccgcctctgc ctctgagcta ttccagaagt 6060
agtgaggagg cttttttgga ggcctaggct tttgcaaaaa gctcccggga gcttgtatat 6120
ccattttcgg atctgatcag cacgtgttga caattaatca tcggcatagt atatcggcat 6180
agtataatac gacaaggtga ggaactaaac catggccaag ttgaccagtg ccgttccggt 6240
gctcaccgcg cgcgacgtcg ccggagcggt cgagttctgg accgaccggc tcgggttctc 6300
ccgggacttc gtggaggacg acttcgccgg tgtggtccgg gacgacgtga ccctgttcat 6360
cagcgcggtc caggaccagg tggtgccgga caacaccctg gcctgggtgt gggtgcgcgg 6420
cctggacgag ctgtacgccg agtggtcgga ggtcgtgtcc acgaacttcc gggacgcctc 6480
cgggccggcc atgaccgaga tcggcgagca gccgtggggg cgggagttcg ccctgcgcga 6540
cccggccggc aactgcgtgc acttcgtggc cgaggagcag gactgacacg tgctacgaga 6600

```
tttcgattcc accgccgcct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc   6660
cggctggatg atcctccagc gcggggatct catgctggag ttcttcgccc accccaactt   6720
gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa   6780
agcatttttt tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca   6840
tgtctgtata ccgtcgacct ctagctagag cttggcgtaa tcatggtcat agctgtttcc   6900
tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg   6960
taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc   7020
cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg   7080
gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc   7140
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac   7200
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa   7260
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca   7320
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc   7380
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata   7440
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta   7500
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca   7560
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga   7620
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg   7680
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg   7740
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg   7800
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag   7860
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa   7920
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat   7980
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc   8040
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc   8100
atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc   8160
tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc   8220
aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc   8280
catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt   8340
gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc   8400
ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa   8460
aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt   8520
atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg   8580
cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc   8640
```

gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa 8700

agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt 8760

gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt 8820

caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag 8880

ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta 8940

tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 9000

aggggttccg cgcacatttc cccgaaaagt gccacctgac gtc 9043

<210> 6
<211> 9658
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 6
[illegible]cggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420

attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480

atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540

atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660

[illegible]tcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720

aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840

ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900

agtccaggaa tcatgctgga gaagttctgc aactctactt tttggaattc ctcattcctg 960

gacagtccgg aggcagacct gccactttgt tttgagcaaa ctgttctggt gtggattccc 1020

ttgggcttcc tatggctcct ggcccccctgg cagcttctcc acgtgtataa atccaggacc 1080

aagagatcct ctaccaccaa actctatctt gctaagcagg tattcgttgg ttttcttctt 1140

attctagcag ccatagagct ggcccttgta ctcacagaag actctggaca agccacagtc 1200

cctgctgttc gatataccaa tccaagcctc tacctaggca catggctcct ggttttgctg 1260

atccaataca gcagacaatg gtgtgtacag aaaaactcct ggttcctgtc cctattctgg 1320

attctctcga tactctgtgg cactttccaa tttcagactc tgatccggac actcttacag 1380
ggtgacaatt ctaatctagc ctactcctgc ctgttcttca tctcctacgg attccagatc 1440
ctgatcctga tctttttcagc attttcagaa aataatgagt catcaaataa tccatcatcc 1500
atagcttcat tcctgagtag cattacctac agctggtatg acagcatcat tctgaaaggc 1560
tacaagcgtc ctctgacact cgaggatgtc tgggaagttg atgaagagat gaaaaccaag 1620
acattagtga gcaagtttga aacgcacatg aagagagagc tgcagaaagc caggcgggca 1680
ctccagagac ggcaggagaa gagctcccag cagaactctg gagccaggct gcctggcttg 1740
aacaagaatc agagtcaaag ccaagatgcc cttgtcctgg aagatgttga aaagaaaaaa 1800
aagaagtctg ggaccaaaaa agatgttcca aaatcctggt tgatgaaggc tctgttcaaa 1860
actttctaca tggtgctcct gaaatcattc ctactgaagc tagtgaatga catcttcacg 1920
tttgtgagtc ctcagctgct gaaattgctg atctcctttg caagtgaccg tgacacatat 2040
ttgtggattg gatatctctg tgcaatcctc ttattcactg cggctctcat tcagtctttc 2040
[illegible]ccttcagt gttatttcca actgtgcttc aagctgggtg taaaagtacg gacagctatc 2100
atggcttctg tatataagaa ggcattgacc ctatccaact tggccaggaa ggagtacacc 2160
gttggagaaa cagtgaacct gatgtctgtg gatgcccaga agctcatgga tgtgaccaac 2220
ttcatgcaca tgctgtggtc aagtgttcta cagattgtct tatctatctt cttcctatgg 2280
agagagttgg gaccctcagt cttagcaggt gttggggtga tggtgcttgt aatcccaatt 2340
aatgcgatac tgtccaccaa gagtaagacc attcaggtca aaaatatgaa gaataaagac 2400
aaacgtttaa agatcatgaa tgagattctt agtggaatca agatcctgaa atattttgcc 2460
tgggaacctt cattcagaga ccaagtacaa aacctccgga agaaagagct caagaacctg 2520
ctggccttta gtcaactaca gtgtgtagta atattcgtct tccagttaac tccagtcctg 2580
gtatctgtgg tcacattttc tgtttatgtc ctggtggata gcaacaatat tttggatgca 2640
caaaaggcct tcacctccat taccctcttc aatatcctgc gctttcccct gagcatgctt 2700
[illegible]catgatga tctcctccat gctccaggcc agtgtttcca cagagcggct agagaagtac 2760
ttgggagggg atgacttgga cacatctgcc attcgacatg actgcaattt tgacaaagcc 2820
atgcagtttt ctgaggcctc ctttacctgg gaacatgatt cggaagccac agtccgagat 2880
gtgaacctgg acattatggc aggccaactt gtggctgtga taggccctgt cggctctggg 2940
aaatcctcct tgatatcagc catgctggga gaaatggaaa atgtccacgg gcacatcacc 3000
atcaagggca ccactgccta tgtcccacag cagtcctgga ttcagaatgg caccataaag 3060
gacaacatcc tttttggaac agagtttaat gaaaagaggt accagcaagt actggaggcc 3120
tgtgctctcc tcccagactt ggaaatgctg cctggaggag atttggctga gattggagag 3180
aagggtataa atcttagtgg gggtcagaag cagcggatca gcctggccag agctacctac 3240
caaaatttag acatctatct tctagatgac cccctgtctg cagtggatgc tcatgtagga 3300
aaacatattt ttaataaggt cttgggcccc aatggcctgt tgaaaggcaa gactcgactc 3360

ttggttacac atagcatgca ctttcttcct caagtggatg agattgtagt tctggggaat 3420
ggaacaattg tagagaaagg atcctacagt gctctcctgg ccaaaaaagg agagtttgct 3480
aagaatctga agacatttct aagacataca ggccctgaag aggaagccac agtccatgat 3540
ggcagtgaag aagaagacga tgactatggg ctgatatcca gtgtggaaga gatccccgaa 3600
gatgcagcct ccataaccat gagaagagag aacagctttc gtcgaacact tagccgcagt 3660
tctaggtcca atggcaggca tctgaagtcc ctgagaaact ccttgaaaac tcggaatgtg 3720
aatagcctga aggaagacga agaactagtg aaaggacaaa aactaattaa gaaggaattc 3780
atagaaactg gaaaggtgaa gttctccatc tacctggagt acctacaagc aataggattg 3840
ttttcgatat tcttcatcat ccttgcgttt gtgatgaatt ctgtggcttt tattggatcc 3900
aacctctggc tcagtgcttg gaccagtgac tctaaaatct tcaatagcac cgactatcca 3960
gcatctcaga gggacatgag agttggagtc tacggagctc tgggattagc ccaaggtata 4020
tttgtgttca tagcacattt ctggagtgcc tttggtttcg tccatgcatc aaatatcttg 4080
[illegible]tcaagcaac tgctgaacaa tatccttcga gcacctatga gatttttga cacaacaccc 4140
acaggccgga ttgtgaacag gtttgccggc gatatttcca cagtggatga caccctgcct 4200
cagtccttgc gcagctggat tacatgcttc ctggggataa tcagcaccct tgtcatgatc 4260
tgcatggcca ctcctgtctt caccatcatc gtcattcctc ttggcattat ttatgtatct 4320
gttcagatgt tttatgtgtc tacctcccgc cagctgaggc gtctggactc tgtcaccagg 4380
tccccaatct actctcactt cagcgagacc gtatcaggtt tgccagttat ccgtgccttt 4440
gagcaccagc agcgatttct gaaacacaat gaggtgagga ttgacaccaa ccagaaatgt 4500
gtcttttcct ggatcacctc caacaggtgg cttgcaattc gcctggagct ggttgggaac 4560
ctgactgtct tcttttcagc cttgatgatg gttatttata gagataccct aagtggggac 4620
actgttggct ttgttctgtc caatgcactc aatatcacac aaaccctgaa ctggctggtg 4680
aggatgacat cagaaataga gaccaacatt gtggctgttg agcgaataac tgagtacaca 4740
[illegible]agtggaaa atgaggcacc ctgggtgact gataagaggc ctccgccaga ttggcccagc 4800
aaaggcaaga tccagtttaa caactaccaa gtgcggtacc gacctgagct ggatctggtc 4860
ctcagaggga tcacttgtga cattggtagc atggagaaga ttggtgtggt gggcaggaca 4920
ggagctggaa agtcatccct cacaaactgc ctcttcagaa tcttagaggc tgccggtggt 4980
cagattatca ttgatggagt agatattgct tccattgggc tccacgacct ccgagagaag 5040
ctgaccatca tcccccagga ccccatcctg ttctctggaa gcctgaggat gaatctcgac 5100
cctttcaaca actactcaga tgaggagatt tggaaggcct tggagctggc tcacctcaag 5160
tcttttgtgg ccagcctgca acttgggtta tcccacgaag tgacagaggc tggtggcaac 5220
ctgagcatag gccagaggca gctgctgtgc ctgggcaggg ctctgcttcg gaaatccaag 5280
atcctggtcc tggatgaggc cactgctgcg gtggatctag agacagacaa cctcattcag 5340
acgaccatcc aaaacgagtt cgcccactgc acagtgatca ccatcgccca caggctgcac 5400

accatcatgg acagtgacaa ggtaatggtc ctagacaacg ggaagattat agagtgcggc 5460
agccctgaag aactgctaca aatccctgga cccttttact ttatggctaa ggaagctggc 5520
attgagaatg tgaacagcac aaaattctag cttaagcttg gtaccgagct cggatccact 5580
agtccagtgt ggtggaattc tgcagatatc cagcacagtg gcggccgctc gagtctagag 5640
ggcccgttta aacccgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg 5700
tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct 5760
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg 5820
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg 5880
cggtgggctc tatggcttct gaggcggaaa gaaccagctg gggctctagg gggtatcccc 5940
acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg 6000
ctacacttgc cagcgcccta gcgcccgctc ctttcgcttt cttcccttcc tttctcgcca 6060
cgttcgccgg ctttccccgt caagctctaa atcgggggcat ccctttaggg ttccgattta 6120
gtgctttacg gcacctcgac cccaaaaaac ttgattaggg tgatggttca cgtagtgggc 6180
catcgccctg atagacggtt tttcgccctt tgacgttgga gtccacgttc tttaatagtg 6240
gactcttgtt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat 6300
aagggatttt ggggatttcg gcctattggt taaaaaatga gctgatttaa caaaaattta 6360
acgcgaatta attctgtgga atgtgtgtca gttagggtgt ggaaagtccc caggctcccc 6420
aggcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccagg tgtggaaagt 6480
ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 6540
tagtcccgcc cctaactccg cccatcccgc ccctaactcc gcccagttcc gcccattctc 6600
cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tctgcctctg 6660
agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaaagctcc 6720
cgggagcttg tatatccatt ttcggatctg atcagcacgt gttgacaatt aatcatcggc 6780
atagtatatc ggcatagtat aatacgacaa ggtgaggaac taaaccatgg ccaagttgac 6840
cagtgccgtt ccggtgctca ccgcgcgcga cgtcgccgga gcggtcgagt tctggaccga 6900
ccggctcggg ttctcccggg acttcgtgga ggacgacttc gccggtgtgg tccgggacga 6960
cgtgaccctg ttcatcagcg cggtccagga ccaggtggtg ccggacaaca ccctggcctg 7020
ggtgtgggtg cgcggcctgg acgagctgta cgccgagtgg tcggaggtcg tgtccacgaa 7080
cttccgggac gcctccgggc cggccatgac cgagatcggc gagcagccgt gggggcggga 7140
gttcgccctg cgcgacccgg ccggcaactg cgtgcacttc gtggccgagg agcaggactg 7200
acacgtgcta cgagatttcg attccaccgc cgccttctat gaaaggttgg gcttcggaat 7260
cgttttccgg gacgccggct ggatgatcct ccagcgcggg gatctcatgc tggagttctt 7320
cgcccacccc aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac 7380
aaatttcaca aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat 7440

caatgtatct tatcatgtct gtataccgtc gacctctagc tagagcttgg cgtaatcatg 7500
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc 7560
cggaagcata aagtgtaaag cctggggtgc ctaatgagtg agctaactca cattaattgc 7620
gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat 7680
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac 7740
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt 7800
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 7860
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc 7920
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 7980
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 8040
gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcaatg 8100
ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca 8160
cgaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa 8220
cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 8280
gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag 8340
aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 8400
tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca 8460
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc 8520
tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag 8580
gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 8640
tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 8700
ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg 8760
ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc 8820
tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc 8880
aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc 8940
gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc 9000
gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc 9060
ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa 9120
gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat 9180
gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata 9240
gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca 9300
tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag 9360
gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc 9420
agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc 9480

aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata 9540
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 9600
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtc 9658